Construction of Network Error Correction Codes in Packet Networks

Xuan Guang, Fang-Wei Fu, and Zhen Zhang, *Fellow, IEEE*

***Abstract*—Recently, network error correction coding (NEC) has been studied extensively. Several bounds in classical coding theory have been extended to network error correction coding, especially the Singleton bound. In this paper, following the research line using the extended global encoding kernels proposed in [12], the refined Singleton bound of NEC can be proved more explicitly. Moreover, we give a constructive proof of the attainability of this bound and indicate that the required field size for the existence of network maximum distance separable (MDS) codes can become smaller further. By this proof, an algorithm is proposed to construct general linear network error correction codes including the linear network error correction MDS codes. Finally, we study the error correction capability of random linear network error correction coding. Motivated partly by the performance analysis of random linear network coding [6], we evaluate the different failure probabilities defined in this paper in order to analyze the performance of random linear network error correction coding. Several upper bounds on these probabilities are obtained and they show that these probabilities will approach to zero as the size of the base field goes to infinity. Using these upper bounds, we slightly improve on the probability mass function of the minimum distance of random linear network error correction codes in [7], as well as the upper bound on the field size required for the existence of linear network error correction codes with degradation at most d.**

***Index Terms*—Network coding, network error correction coding, the refined Singleton bound, maximum distance separable (MDS) code, random linear network error correction coding, the extended global encoding kernels, network error correction code construction.**

I. INTRODUCTION

NETWORK coding was first introduced by Yeung and Zhang in [1] and then was profoundly developed by Ahlswede *et al.* [2]. In the latter paper [2], the authors showed that by network coding in network communication, the source node can multicast the information to all sink nodes at the theoretically maximum rate as the alphabet size approaches infinity, where the theoretically maximum rate is the smallest minimum cut capacity between the source node and any sink node. Li *et al.* [3] indicated that linear network coding with finite alphabet size is sufficient for multicast. In [4], Koetter and Médard presented an algebraic characterization for network coding. Although network coding can achieve the higher information rate than classical routing, Jaggi *et al.* [5] still proposed a deterministic polynomial-time algorithm for constructing a linear network code. Random linear network coding was originally introduced by Ho *et al.* [6], and the authors analyzed the performance of random linear network coding by studying the failure probabilities of the codes. Balli, Yan, and Zhang [7] improved on the upper bounds on these failure probabilities and then studied the asymptotic behavior of the failure probability as the field size goes to infinity. Following [7], Guang and Fu [8] gave some tight or asymptotically tight bounds for two kinds of failure probabilities and also gave the specific network structures in the worst cases.

Network coding has been extensively studied for several years under the assumption that channels of networks are error-free. Unfortunately, all kinds of errors may occur in network communication such as random errors, erasure errors (packet losses), errors in headers and so on. In order to deal with such problems, network error correction coding (NEC) was studied recently. The original idea of network error correction coding was proposed by Yeung and Cai in their conference paper [9] and developed in their recent journal papers [10] [11]. In the latter two papers, the concept of network error correction codes was introduced as a generalization of the classical error correction codes. They also extended some important bounds from classical error correction codes to network error correction codes, such as the Singleton bound, the Hamming bound, and the Gilbert-Varshamov bound. Although the Singleton bound has been given in Cai and Yeung [10], Zhang [12] and Yang *et al.* [13] [14] presented the refined Singleton bound independently by using the different methods. Yang *et al.* [15] [14] developed a framework for characterizing error correction/detection capabilities of network error correction codes. They defined different minimum distances to measure error correction and error detection capabilities, respectively. It followed an interesting discovery that, for nonlinear network error correction codes, the number of the correctable errors can be more than half of the number of the detectable errors. In [12], Zhang defined the minimum distance of linear network error correction codes and introduced the concept of extended global encoding kernels. Using this concept, Zhang proposed linear network error correction codes in packet networks. Besides coherent networks, this scheme is also suitable to non-coherent networks by recording the extended global encoding kernels in the headers of the packets. Moreover, the extended global encoding kernels are used to form the decoding matrices at sink nodes. As well as in [16], the decoding principles and decoding beyond the error correction capability were studied. The authors further presented several decoding algorithms and

This research is supported in part by the National Natural Science Foundation of China under the Grants 60872025 and 10990011.

X. Guang is with the Chern Institute of Mathematics, Nankai University, Tianjin 300071, P.R. China. Email: xuanguang@mail.nankai.edu.cn.

F.-W. Fu is with the Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, P.R. China. Email: fwfu@nankai.edu.cn.

Z. Zhang is with the Communication Sciences Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2565 USA. Email: zhzhang@usc.edu.

analyzed their performance. In addition, Balli, Yan, and Zhang [7] studied the error correction capability of random linear network error correction codes. They gave the probability mass function of the minimum distance of random linear network error correction codes. For the existence of a network error correction code with degradation, the upper bound on the required field size was proposed.

In [17], Koetter and Kschischang formulated a different framework for network error correction coding. In their approach, the source message is represented by a subspace of a fixed vector space and a basis of the subspace is injected into the network. This type of network error correction codes is called subspace codes.

In this paper, we follow the research line using the extended global encoding kernels introduced by Zhang in [18] [12]. We reprove the refined Singleton bound of the network error correction codes more explicitly by using the concept of the extended global encoding kernels. Similar to the Singleton bound in classical coding theory, the refined Singleton bound is also tight and those linear network error correction codes achieving this bound with equality are called linear network error correction maximum distance separable (MDS) codes, or network MDS codes for short. For network MDS codes, Zhang [12] gave an existence proof by an algebraic method. In this paper, we present a constructive proof of the attainability of the refined Singleton bound, and indicate that the required field size for the existence of network MDS codes can become smaller (in some cases much smaller) than the known results. Moreover, by this proof, we design an algorithm for constructing general linear network error correction codes, in particular, network MDS codes.

Matsumoto [19] and Yang *et al.* [13] also proposed the algorithms for constructing network MDS codes. The algorithm of Yang *et al.* designs the codebook and the local encoding kernels separately. On the contrary, Matsumoto's algorithm and our algorithm design them together. As noted above, the required field size of our algorithm is smaller. Moreover, compared with Matsumoto's algorithm, our algorithm needs less storages at each sink node. For the decoding, as mentioned in [19], the decoding of Matsumoto's algorithm requires exhaustive search by each sink node for all possible information from the source and all possible errors, and our algorithm can make use of the better and faster decoding algorithms proposed by Zhang, Yan, and Balli in a series of papers [18], [12], and [16] such as the brute force decoding algorithm and the fast decoding algorithm. For the case of decoding network error correction codes beyond the error correction capability in packet networks [16], our algorithm has more advantages because of the use of extended global encoding kernels. We further study the error correction capability of random linear network error correction coding, and analyze the failure probabilities of constructing network MDS codes and general network error correction codes by using random method, as well as the probability mass function of the minimum distance and the required field size.

This paper is divided into 6 sections. In the next section, we introduce the basic notation and definitions in linear network coding and linear network error correction coding, and give some propositions needed in this paper. In Section III, we reprove the refine Singleton bound by using the concept of the extended global encoding kernels, and propose a constructive proof to show the attainability of the refined Singleton bound of NEC. Consequently, we also indicate that the required field size for the existence of network MDS codes can become smaller than the known results. Section IV is devoted to the algorithm for constructing general linear network error correction codes, including network MDS codes. In Section V, we analyze the performance of random linear network error correction codes . The last section summarizes the works done in this paper.

II. Basic Notation and Definitions

In this paper, we follow [12] in its notation and terminology. A communication network is defined as a finite acyclic directed graph $G = (V, E)$, where the vertex set V stands for the set of nodes and the edge set E represents the set of communication channels of the network. The node set V consists of three disjoint subsets S, T, and J, where S is the set of source nodes, T is the set of sink nodes, and $J = V - S - T$ is the set of internal nodes. Furthermore, a direct edge $e = (i, j) \in E$ represents a channel leading from node i to node j. Node i is called the tail of e and node j is called the head of e, written as $i = tail(e)$, $j = head(e)$, respectively. Correspondingly, the channel e is called an outgoing channel of i and an incoming channel of j. For a node i, define $Out(i) = \{e \in E : e \text{ is an outgoing channel of } i\}$, $In(i) = \{e \in E : e \text{ is an incoming channel of } i\}$. In a communication network, if a sequence of channels $(e_1, e_2, \cdots, e_m)$ satisfies $tail(e_1) = i$, $head(e_m) = j$, and $tail(e_{k+1}) = head(e_k)$ for $k = 1, 2, \cdots, m-1$, then we call the sequence $(e_1, e_2, \cdots, e_m)$ a path from node i to node j, or equivalently, a path from channel e_1 to node j. For each channel $e \in E$, there exists a positive number R_e called the capacity of e. We allow the multiple channels between two nodes and assume reasonably that the capacity of any channel is 1 per unit time. This means that one field symbol can be transmitted over a channel in one unit time. A cut between node i and node j is a set of channels whose removal disconnects i from j. For unit capacity channels, the capacity of a cut can be regarded as the number of channels in the cut, and the minimum of all capacities of cuts between i and j is called the minimum cut capacity between node i and node j. A cut between node i and node j is called a minimum cut if its capacity achieves the minimum cut capacity between i and j. Note that there may exist several minimum cuts between i and j, but the minimum cut capacity between them is determined. The source nodes generate messages and transmit them to all sink nodes over the network by network coding. In the present paper, we consider single source networks, i.e., $|S| = 1$, and the unique source node is denoted by s. The source node s has no incoming channels and any sink node has no outgoing channels, but we use the concept of imaginary incoming channels of the source node s and assume that these imaginary incoming channels provide the source messages to s. Let the information rate be w symbols per unit time. Then the source node has w imaginary incoming channels $d'_1, d'_2, \cdots, d'_w$ and let $In(s) = \{d'_1, d'_2, \cdots, d'_w\}$. The source messages are w symbols $\underline{\mathbf{X}} = (X_1, X_2, \cdots, X_w)$ arranged in a row vector where each X_i is an element of base field $\mathcal{F}$. Assume that they are transmitted to the source node s through the w imaginary channels in $In(s)$. By using network coding, source messages are multicast to and decoded at each sink node.

At each node $i \in V - T$, there is an $|In(i)| \times |Out(i)|$ matrix $K_i = (k_{d,e})_{d \in In(i), e \in Out(i)}$ called the local encoding kernel at i, where $k_{d,e} \in \mathcal{F}$ is called the local encoding coefficient for the adjacent pair of channels (d, e). Denote by U_e the message transmitted over the channel e. At the source node s, assume that the message transmitted over the ith imaginary channel is the ith source message, i.e., $U_{d'_i} = X_i$. In general, the message U_e is calculated by the formula $U_e = \sum_{d \in In(tail(e))} k_{d,e} U_d$. As we know from [20] [21], the global encoding kernel of a channel e is a w-dimensional column vector f_e over the base field $\mathcal{F}$ satisfying $U_e = \underline{\mathbf{X}} \cdot f_e$. The global encoding kernels can be determined by the local encoding kernels.

In the case that there is an error in a channel e, the output of the channel is $\tilde{U}_e = U_e + Z_e$, where U_e is the message that should be transmitted over the channel e and $Z_e \in \mathcal{F}$ is the error occurred in e. We treat Z_e as a message called *error message*. To explain the approach, the extended network was introduced in [12] as follows. In the original network $G = (V, E)$, for each channel $e \in E$, an imaginary channel e' is introduced, which is connected to the tail of e to provide error message. This network $\tilde{G} = (\tilde{V}, \tilde{E})$ with imaginary channels is called the extended network, where $\tilde{V} = V$ and $\tilde{E} = E \cup E' \cup \{d'_1, d'_2, \cdots, d'_w\}$ with $E' = \{e' : e \in E\}$. Obviously, $|E'| = |E|$. Then a linear network code for the original network can be extended to a linear network code for the extended network by letting $k_{e',e} = 1$ and $k_{e',d} = 0$ for all $d \in E \backslash \{e\}$. For each internal node i in the extended network, note that $In(i)$ only includes the real incoming channels of i, that is, the imaginary channels e' corresponding to $e \in Out(i)$ are not in $In(i)$. But for the source node s, we still define $In(s) = \{d'_1, d'_2, \cdots, d'_w\}$. In order to distinguish two different types of imaginary channels, we call d'_i $(1 \leq i \leq w)$ the imaginary message channels and e' for $e \in E$ the imaginary error channels. We can also define global encoding kernel $\tilde{f}_e$ for each $e \in \tilde{E}$ in the extended network. It is a $(w + |E|)$-dimensional column vector and the entries can be indexed by the elements of $In(s) \cup E$. For imaginary message channels d'_i $(1 \leq i \leq w)$ and imaginary error channels $e' \in E'$, let $\tilde{f}_{d'_i} = 1_{d'_i}$, $\tilde{f}_{e'} = 1_e$, where 1_d is a $(w + |E|)$-dimensional column vector which is the indicator function of $d \in In(s) \cup E$. For other global encoding kernels $\tilde{f}_e, e \in E$, we have recursive formulae:

$$\tilde{f}_e = \sum_{d \in In(tail(e))} k_{d,e} \tilde{f}_d + 1_e.$$

We call $\tilde{f}_e$ the extended global encoding kernel of the channel e $(e \in E)$ for the original network. Furthermore, similar to the Koetter-Médard Formula [4], there also exists a formula [12]:

$$(\tilde{f}_e : \ e \in E) = \begin{pmatrix} B \\ I \end{pmatrix} (I - F)^{-1},$$

where $B = (k_{d,e})_{d \in In(s), e \in E}$ is a $w \times |E|$ matrix with $k_{d,e} = 0$ for $e \notin Out(s)$ and $k_{d,e}$ being the local encoding coefficient for $e \in Out(s)$, the system transfer matrix $F = (k_{d,e})_{d \in E, e \in E}$ is an $|E| \times |E|$ matrix with $k_{d,e}$ being the local encoding coefficient for $head(d) = tail(e)$ and $k_{d,e} = 0$ for $head(d) \neq tail(e)$, and I is an $|E| \times |E|$ identity matrix.

Let $\underline{\mathbf{Z}} = (Z_e : \ e \in E)$ be an $|E|$-dimensional row vector with $Z_e \in \mathcal{F}$ for all $e \in E$. Then $\underline{\mathbf{Z}}$ is called the error message vector. An error pattern ρ is regarded as a set of channels in which errors occur. We call that an error message vector $\underline{\mathbf{Z}}$ matches an error pattern ρ, if $Z_e = 0$ for all $e \in E \backslash \rho$.

For a channel $e \in E$, if there is no error in it, then

$$\tilde{U}_e = (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot \tilde{f}_e = (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot (\tilde{f}_e - 1_e) = U_e.$$

If there is an error $Z_e \neq 0$ in channel e, then

$$\begin{aligned} \tilde{U}_e &= U_e + Z_e = (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot (\tilde{f}_e - 1_e) + Z_e \\ &= (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot (\tilde{f}_e - 1_e) + (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot 1_e = (\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \cdot \tilde{f}_e. \end{aligned}$$

At a sink node t, the messages $\{\tilde{U}_e : \ e \in In(t)\}$ and the extended global encoding kernels $\{\tilde{f}_e : \ e \in In(t)\}$ are available. For all messages including information messages and error messages, if they are considered as column vectors, then the above discussions describe linear network error correction coding in packet networks.

First, we need some notation and definitions which either are quoted directly or are extended from Zhang [12].

Definition 1 ([12, Definition 1]): The matrix

$$\tilde{F}_t = (\tilde{f}_e : \ e \in In(t))$$

is called the decoding matrix at a sink node $t \in T$. Let

$$\tilde{A}_t = (\tilde{U}_e : \ e \in In(t)).$$

The equation

$$(\underline{\mathbf{X}}, \underline{\mathbf{Z}}) \tilde{F}_t = \tilde{A}_t$$

is called the decoding equation at a sink node t.

Definition 2: For an error pattern ρ and extended global encoding kernels $\tilde{f}_e, \ e \in E$,

- $\tilde{f}_e^\rho$ is a $(w + |\rho|)$-dimensional column vector obtained from $\tilde{f}_e = (\tilde{f}_e(d) : d \in In(s) \cup E)$ by removing all entries $\tilde{f}_e(d), \ d \notin In(s) \cup \rho$, and $\tilde{f}_e^\rho$ is called the extended global encoding kernel of channel e restricted to the error pattern ρ.
- f_e^ρ is a $(w + |E|)$-dimensional column vector obtained from $\tilde{f}_e = (\tilde{f}_e(d) : d \in In(s) \cup E)$ by replacing all entries $\tilde{f}_e(d), \ d \notin In(s) \cup \rho$ by 0, and f_e^ρ is also called the extended global encoding kernel of channel e restricted to the error pattern ρ.
- $f_e^{\rho^c}$ is a $(w + |E|)$-dimensional column vector obtained from $\tilde{f}_e = (\tilde{f}_e(d) : d \in In(s) \cup E)$ by replacing all entries $\tilde{f}_e(d), \ d \in In(s) \cup \rho$ by 0.

Note that $f_e^\rho + f_e^{\rho^c} = \tilde{f}_e$.

Definition 3 ([12, Defintion 3]): Define

$$\Delta(t, \rho) = \{(\underline{\mathbf{0}}, \underline{\mathbf{Z}}) \tilde{F}_t : \ \text{all } \underline{\mathbf{Z}} \text{ matching the error pattern } \rho\}$$

where $\underline{\mathbf{0}}$ is a w-dimensional zero row vector, and $\underline{\mathbf{Z}}$ is an $|E|$-dimensional row vector matching the error pattern ρ; and

$$\Phi(t) = \{(\underline{\mathbf{X}}, \underline{\mathbf{0}}) \tilde{F}_t : \ \underline{\mathbf{X}} \in \mathcal{F}^w\}.$$

We call $\Delta(t,\rho)$ the error space of error pattern ρ and $\Phi(t)$ the message space.

Let L be a collection of vectors in a linear space. $\langle L\rangle$ represents the subspace spanned by the vectors in L. In fact, if we use $row_t(d),\ d\in In(s)\cup E$ to denote the row vectors of the decoding matrix $\tilde{F}_t$, then $\Delta(t,\rho)=\langle\{row_t(d):\ d\in\rho\}\rangle$ and $\Phi(t)=\langle\{row_t(d):\ d\in In(s)\}\rangle$.

Definition 4 ([12, Definition 4]): We say that an error pattern ρ_1 is dominated by another error pattern ρ_2 with respect to a sink node t if $\Delta(t,\rho_1)\subseteq\Delta(t,\rho_2)$ for any linear network code. This relation is denoted by $\rho_1\prec_t\rho_2$.

Definition 5 ([12, Definition 5]): The rank of an error pattern ρ with respect to a sink node t is defined by

$$rank_t(\rho)=\min\{|\rho'|:\ \rho\prec_t\rho'\}.$$

In order to understand the concept of rank of an error pattern better, we give the following proposition. This proposition is a slight and necessary modification of [12, Lemma 1].

Proposition 1: For an error pattern ρ, introduce a source node s_ρ. Let $\rho=\{e_1,e_2,\cdots,e_l\}$ where $e_j\in In(i_j)$ for $1\leq j\leq l$ and define $e'_j=(s_\rho,i_j)$. Replace each e_j by e'_j on the network, that is, add $e'_1,e'_2,\cdots,e'_l$ on the network and delete $e_1,e_2,\cdots,e_l$ from the network. Then the rank of the error pattern ρ with respect to a sink node t is equal to the minimum cut capacity between s_ρ and t.

Proof: It is similar to the proof in [12], and, therefore, omitted. ■

Definition 6 ([12, Definition 6]): A linear network error correction code is called a regular code if for any $t\in T$, $\dim(\Phi(t))=w$.

Definition 7 ([12, Definition 7]): The minimum distance of a regular network error correction code at a sink node t is defined by

$$d_{\min}^{(t)}=\min\{rank_t(\rho):\ \dim(\Delta(t,\rho)\cap\Phi(t))>0\}.$$

For the minimum distance above, we give the following proposition.

Proposition 2: For the minimum distance of a regular network error correction code at a sink node t, there exist the following equalities:

$$\begin{aligned}d_{\min}^{(t)}&=\min\{rank_t(\rho):\ \Delta(t,\rho)\cap\Phi(t)\neq\{\underline{0}\}\} &(1)\\ &=\min\{|\rho|:\ \Delta(t,\rho)\cap\Phi(t)\neq\{\underline{0}\}\} &(2)\\ &=\min\{\dim(\Delta(t,\rho)):\ \Delta(t,\rho)\cap\Phi(t)\neq\{\underline{0}\}\}. &(3)\end{aligned}$$

Proof: We define the set of error patterns $\Pi=\{\rho:\ \Delta(t,\rho)\cap\Phi(t)\neq\{\underline{0}\}\}$. Then one has

$$(1)=\min_{\rho\in\Pi}rank_t(\rho),\ (2)=\min_{\rho\in\Pi}|\rho|,\ (3)=\min_{\rho\in\Pi}\dim(\Delta(t,\rho)).$$

Since $\dim(\Delta(t,\rho))\leq rank_t(\rho)\leq|\rho|$ for any error pattern $\rho\subseteq E$, it follows that

$$\min_{\rho\in\Pi}\dim(\Delta(t,\rho))\leq\min_{\rho\in\Pi}rank_t(\rho)\leq\min_{\rho\in\Pi}|\rho|.$$

In view of the inequalities above, it is enough to prove $\min_{\rho\in\Pi}|\rho|\leq\min_{\rho\in\Pi}\dim(\Delta(t,\rho))$. Let $\rho'\in\Pi$ be an error pattern satisfying

$$\dim(\Delta(t,\rho'))=\min_{\rho\in\Pi}\dim(\Delta(t,\rho)).$$

Assume that $\rho'=\{e_1,e_2,\cdots,e_l\}$, which means $\Delta(t,\rho')=\langle\{row_t(e_i):\ 1\leq i\leq l\}\rangle$. For $\{row_t(e_i):\ 1\leq i\leq l\}$, let its maximum independent vector set be $\{row_t(e_{i_j}):1\leq j\leq m\}$, where $m=\dim(\Delta(t,\rho'))\leq l$. Set $\rho_1=\{e_{i_j}:1\leq j\leq m\}$. This implies that

$$|\rho_1|=\dim(\Delta(t,\rho_1))=\dim(\Delta(t,\rho'))$$

and

$$\Delta(t,\rho_1)\cap\Phi(t)=\Delta(t,\rho')\cap\Phi(t)\neq\{\underline{0}\}.$$

Therefore,

$$\min_{\rho\in\Pi}|\rho|\leq|\rho_1|=\dim(\Delta(t,\rho'))=\min_{\rho\in\Pi}\dim(\Delta(t,\rho)).$$

The proof is completed. ■

In this paper, we always use w to denote the information rate and C_t to denote the minimum cut capacity between the unique source node s and sink node t, and define $\delta_t=C_t-w$ which is called the redundancy of sink node t.

III. The Refined Singleton Bound of NEC and The Network MDS Codes

By using the concept of the extended global encoding kernels, we can reprove the refined Singleton bound of NEC. First, we give the following lemma.

Lemma 1: For a regular linear network error correction code, let a channel set $\{e_1,e_2,\cdots,e_{C_t}\}$ be a minimum cut between s and t with an upstream-to-downstream order $e_1\prec e_2\prec\cdots\prec e_{C_t}$ and let an error pattern $\rho=\{e_w,e_{w+1},\cdots,e_{C_t}\}$. Then $\Phi(t)\cap\Delta(t,\rho)\neq\{\underline{0}\}$.

Proof: Let $\underline{\mathbf{X}}$ and $\underline{\mathbf{Z}}$ represent the source message vector and the error message vector, respectively. Then, for each channel $e\in E$, we have $\tilde{U}_e=(\underline{\mathbf{X}},\underline{\mathbf{Z}})\cdot\tilde{f}_e$, where $\tilde{U}_e$ is the output of e. Let $\tilde{U}_{e_1}=\tilde{U}_{e_2}=\cdots=\tilde{U}_{e_{w-1}}=0$. Since $\mathrm{Rank}((\tilde{f}_{e_1}\ \tilde{f}_{e_2}\ \cdots\ \tilde{f}_{e_{w-1}}))$ is at most $(w-1)$, there exists a nonzero message vector $\underline{\mathbf{X}}_1$ and an error message vector $\underline{\mathbf{Z}}_1=\underline{\mathbf{0}}$ such that

$$\begin{aligned}&(\underline{\mathbf{X}}_1,\underline{\mathbf{Z}}_1)\cdot(\tilde{f}_{e_1}\ \tilde{f}_{e_2}\ \cdots\ \tilde{f}_{e_{w-1}})\\ =&(\underline{\mathbf{X}}_1,\underline{\mathbf{0}})\cdot(\tilde{f}_{e_1}\ \tilde{f}_{e_2}\ \cdots\ \tilde{f}_{e_{w-1}})\\ =&(\tilde{U}_{e_1}\ \tilde{U}_{e_2}\ \cdots\ \tilde{U}_{e_{w-1}})=\underline{\mathbf{0}}.\end{aligned}$$

Moreover, as this code is regular, this implies

$$(\underline{\mathbf{X}}_1,\underline{\mathbf{0}})\cdot(\tilde{f}_{e_1}\ \tilde{f}_{e_2}\ \cdots\ \tilde{f}_{e_{C_t}})=(\tilde{U}_{e_1}\ \tilde{U}_{e_2}\ \cdots\ \tilde{U}_{e_{C_t}})\neq\underline{\mathbf{0}}.$$

Assume the contrary, i.e., $(\tilde{U}_{e_1}\ \tilde{U}_{e_2}\ \cdots\ \tilde{U}_{e_{C_t}})=\underline{\mathbf{0}}$. And note that $\{e_1,e_2,\cdots,e_{C_t}\}$ is a minimum cut between s and t and $\underline{\mathbf{Z}}_1=\underline{\mathbf{0}}$. It follows that

$$\tilde{A}_t=(\tilde{U}_e:e\in In(t))=\underline{\mathbf{0}},$$

which implies that $(\underline{\mathbf{X}}_1,\underline{\mathbf{0}})\tilde{F}_t=\underline{\mathbf{0}}$ from the decoding equation $(\underline{\mathbf{X}}_1,\underline{\mathbf{Z}}_1)\tilde{F}_t=\tilde{A}_t$. Therefore, the equality $\underline{\mathbf{X}}_1=\underline{\mathbf{0}}$ follows from $\dim(\Phi(t))=w$ because the linear network error correction code considered is regular. This contradicts our assumption $\underline{\mathbf{X}}_1\neq\underline{\mathbf{0}}$.

On the other hand, there exists another source message vector $\underline{\mathbf{X}}_2=\underline{\mathbf{0}}$ and another error message vector $\underline{\mathbf{Z}}_2$ matching the error pattern $\rho=\{e_w,e_{w+1},\cdots,e_{C_t}\}$, such that

$$(\underline{\mathbf{X}}_2,\underline{\mathbf{Z}}_2)\cdot(\tilde{f}_{e_1}\ \tilde{f}_{e_2}\ \cdots\ \tilde{f}_{e_{C_t}})=(\tilde{U}_{e_1}\ \tilde{U}_{e_2}\ \cdots\ \tilde{U}_{e_{C_t}}).$$

And note that $\underline{\mathbf{Z}}_2 \neq \underline{\mathbf{0}}$ because $(\tilde{U}_{e_1}\ \tilde{U}_{e_2}\ \cdots\ \tilde{U}_{e_{C_t}}) \neq \underline{\mathbf{0}}$. In fact, since $e_w \prec e_{w+1} \prec \cdots \prec e_{C_t}$, for $e \in \rho$, we can set sequentially:

$$Z_e = \tilde{U}_e - \sum_{d\in In(tail(e))} k_{d,e}\tilde{U}'_d,$$

where $\tilde{U}'_d$ is the output of channel d in this case.

Therefore, it follows that

$$(\underline{\mathbf{X}}_1, \underline{\mathbf{0}}) \cdot \tilde{F}_t = (\underline{\mathbf{0}}, \underline{\mathbf{Z}}_2) \cdot \tilde{F}_t.$$

And note that $\underline{\mathbf{Z}}_2$ matches the error pattern ρ. It is shown that $\Phi(t) \cap \Delta(t,\rho) \neq \{\underline{0}\}$. The lemma is proved. ■

Theorem 2 (The Refined Singleton Bound): Let $d^{(t)}_{\min}$ be the minimum distance of a regular linear network error correction code at a sink node $t \in T$. Then

$$d^{(t)}_{\min} \leq \delta_t + 1.$$

Remark 1: Conventionally, if a regular network error correction code $\mathbf{C}$ satisfies the refined Singleton bound with equality, that is, $d^{(t)}_{\min} = \delta_t + 1$ for each $t \in T$, then this code $\mathbf{C}$ is called network error correction maximum distance separable (MDS) code, or network MDS codes for short.

It is not hard to see that Theorem 2 is an obvious consequence of Proposition 2 and Lemma 1. Now, we give a constructive proof to show that the refined Singleton bound is tight. First, we need the following lemma from [12]. Define $R_t(\delta_t)$ as the set of the error patterns ρ satisfying $|\rho| = rank_t(\rho) = \delta_t$, that is,

$$R_t(\delta_t) = \{\text{error pattern } \rho : \ |\rho| = rank_t(\rho) = \delta_t\}.$$

Lemma 3: For each $t \in T$ and any error pattern $\rho \in R_t(\delta_t)$, there exist $(w + \delta_t)$ channel disjoint paths from either s or ρ to t, and the $(w + \delta_t)$ paths satisfy the properties that

1) there are exactly δ_t paths from ρ to t, and w paths from s to t;
2) these δ_t paths from ρ to t start with the different channels in ρ.

Furthermore, in Lemma 3, assign w imaginary message channels $d'_1, d'_2, \cdots, d'_w$ to the w paths from s to t, and assign δ_t imaginary error channels $e', e \in \rho$ to the δ_t paths from ρ to t, i.e., for each $e \in \rho$, assign e' to the path from e to t. This leads to the following corollary.

Corollary 4: For each $t \in T$ and any error pattern $\rho \in R_t(\delta_t)$, there exist $(w + \delta_t)$ channel disjoint paths from either $In(s) = \{d'_1, d'_2, \cdots, d'_w\}$ or $\rho' = \{e' : e \in \rho\}$ to t, and the $(w + \delta_t)$ paths satisfy the properties that

1) there are exactly δ_t paths from ρ' to t, and w paths from $In(s)$ to t;
2) these δ_t paths from ρ' to t start with the distinct channels in ρ' and for each path, if it starts with $e' \in \rho'$, then it passes through $e \in \rho$.

Theorem 5: If $|\mathcal{F}| \geq \sum_{t\in T} |R_t(\delta_t)|$, then there exist linear network error correction MDS codes, i.e., for all $t \in T$,

$$d^{(t)}_{\min} = \delta_t + 1.$$

Proof: Let $G = \{V, E\}$ be a single source multicast network, where s is the single source, T is the set of sink nodes, $J = V - \{s\} - T$ is the set of internal nodes, and E represents the set of channels in G. Let $\tilde{G} = (\tilde{V}, \tilde{E})$ be the extended network of G. For each $t \in T$ and each $\rho \in R_t(\delta_t)$, $\mathcal{P}_{t,\rho}$ denotes the set of $(w+\delta_t)$ channel disjoint paths satisfying Corollary 4. Denote by $E_{t,\rho}$ the set of all channels on paths in $\mathcal{P}_{t,\rho}$.

Now, we define a dynamic set of channels $CUT_{t,\rho}$ for each $t \in T$ and each $\rho \in R_t(\delta_t)$, and initialize

$$CUT_{t,\rho} = In(s) \cup \rho' = \{d'_1, d'_2, \cdots, d'_w\} \cup \{e' : \ e \in \rho\},$$

where e' is the imaginary error channel corresponding to e. Initialize $\tilde{f}_d = \underline{0}$ for all $d \in E$ and $\tilde{f}_d = 1_d$ for all $d \in In(s) \cup E'$. Naturally, we are interested in $\{\tilde{f}_d : d \in CUT_{t,\rho}\}$.

For any subset $B \subseteq In(s) \cup E' \cup E$, define

$$\tilde{\mathcal{L}}(B) = \langle\{\tilde{f}_e : e \in B\}\rangle, \tilde{\mathcal{L}}^\rho(B) = \langle\{\tilde{f}^\rho_e : e \in B\}\rangle, \text{ and}$$
$$\mathcal{L}^\rho(B) = \langle\{f^\rho_e : e \in B\}\rangle, \mathcal{L}^{\rho^c}(B) = \langle\{f^{\rho^c}_e : e \in B\}\rangle.$$

For $CUT_{t,\rho}$, note that the initial set is $CUT_{t,\rho} = In(s) \cup \rho'$, which means

$$\begin{aligned}\tilde{\mathcal{L}}(CUT_{t,\rho}) &= \langle\{\tilde{f}_d : d \in In(s) \cup \rho'\}\rangle \\ &= \langle\{1_d : d \in In(s) \cup \{e' : e \in \rho\}\}\rangle.\end{aligned}$$

Thus $\left(\tilde{f}^\rho_d : d \in CUT_{t,\rho}\right) = \left(\tilde{f}^\rho_d : d \in In(s) \cup \rho'\right)$ is an identity matrix of size $(w + \delta_t) \times (w + \delta_t)$. That is, $\mathrm{Rank}(\left(\tilde{f}^\rho_d : d \in In(s) \cup \rho'\right)) = w + \delta_t$ or $\dim(\tilde{\mathcal{L}}^\rho(CUT_{t,\rho})) = w + \delta_t$.

Next, we will update $CUT_{t,\rho}$ in the topological order of all nodes until $CUT_{t,\rho} \subseteq In(t)$.

For each $i \in V$, consider all channels $e \in Out(i)$ in arbitrary order. For each $e \in Out(i)$, if $e \notin \cup_{t\in T} \cup_{\rho\in R_t(\delta_t)} E_{t,\rho}$, let $\tilde{f}_e = 1_e$, and all $CUT_{t,\rho}$ remain unchanged. Otherwise $e \in \cup_{t\in T} \cup_{\rho\in R_t(\delta_t)} E_{t,\rho}$, i.e., $e \in E_{t,\rho}$ for some $t \in T$ and $\rho \in R_t(\delta_t)$. In $\mathcal{P}_{t,\rho}$, we use $e(t,\rho)$ to denote the previous channel of e on the path which e locates on. Choose

$$\begin{aligned}&\tilde{g}_e \in \tilde{\mathcal{L}}(In(i) \cup \{e'\}) \backslash && (4)\\ &\cup_{t\in T} \cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^\rho(CUT_{t,\rho} \backslash \{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})],\end{aligned}$$

where the addition "+" represents the sum of two vector spaces. Further, let

$$\tilde{f}_e = \begin{cases} \tilde{g}_e + 1_e & \text{if } \tilde{g}_e(e) = 0, \\ \tilde{g}_e(e)^{-1} \cdot \tilde{g}_e & \text{otherwise.} \end{cases}$$

For those $CUT_{t,\rho}$ satisfying $e \in E_{t,\rho}$, update $CUT_{t,\rho} = \{CUT_{t,\rho} \backslash \{e(t,\rho)\}\} \cup \{e\}$; and for others, $CUT_{t,\rho}$ remain unchanged.

Updating all channels in E by the same method, one can see that all $\tilde{f}_e, e \in E$ are well-defined and, finally, $CUT_{t,\rho} \subseteq In(t)$ for all $t \in T$ and $\rho \in R_t(\delta_t)$.

To complete the proof, we only need to prove the following two conclusions:

1) For each $t \in T$, $d^{(t)}_{\min} = \delta_t + 1$.
2) There exists nonzero column vector $\tilde{g}_e$ satisfying (4).

The proof of 1): We will indicate that all $CUT_{t,\rho}$ satisfy $\dim(\tilde{\mathcal{L}}^\rho(CUT_{t,\rho})) = w+\delta_t$ during the whole updating process by induction.

Assume that all channels before e have been updated and $\dim(\tilde{\mathcal{L}}^{\rho}(CUT_{t,\rho})) = w + \delta_t$ for each $CUT_{t,\rho}$. Now, we take the channel e into account. Since we choose

$$\tilde{g}_e \in \tilde{\mathcal{L}}(In(i) \cup \{e'\}) \backslash \cup_{t\in T} \cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho} \backslash \{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i)\cup\{e'\})],$$

it follows that $\tilde{g}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are linearly independent for any $CUT_{t,\rho}$ with $e \in E_{t,\rho}$. Conversely, suppose that $\tilde{g}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are linearly dependent. This means that g_e^{ρ} is a linear combination of vectors in $\{f_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$. And $g_e^{\rho^c}$ is a linear combination of vectors in $\{f_d^{\rho^c} : d \in In(i) \cup \{e'\}\}$ because of $\tilde{g}_e \in \tilde{\mathcal{L}}(In(i) \cup \{e'\})$. Therefore, $\tilde{g}_e = g_e^{\rho} + g_e^{\rho^c}$ is a linear combination of vectors in

$$\{f_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\} \cup \{f_d^{\rho^c} : d \in In(i) \cup \{e'\}\}.$$

This is a contradiction to the choice of $\tilde{g}_e$.

In the following, we will show that $\tilde{f}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are also linearly independent.

- If $\tilde{g}_e(e) \neq 0$, then, since $\tilde{g}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are linearly independent, $\tilde{f}_e^{\rho} = \tilde{g}_e(e)^{-1} \cdot \tilde{g}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are also linearly independent.
- Otherwise $\tilde{g}_e(e) = 0$. We claim that $e \notin \rho$. Assume the contrary, i.e., $e \in \rho$. Thus $e(t,\rho) = e'$ which means $\tilde{f}_{e(t,\rho)} = 1_e$ and $\tilde{f}_d(e) = 0$ for all $d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}$. Together with $\tilde{g}_e(e) = 0$ and $\dim(\tilde{\mathcal{L}}^{\rho}(CUT_{t,\rho})) = w+\delta_t$, it follows that $\tilde{g}_e^{\rho}$ is a linear combination of vectors in $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$. This implies that $\tilde{g}_e \in \mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})$, which leads to a contradiction. Hence, in view of $e \notin \rho$, one obtains $\tilde{g}_e^{\rho} = \tilde{f}_e^{\rho}$, which implies that $\tilde{f}_e^{\rho}$ and $\{\tilde{f}_d^{\rho} : d \in CUT_{t,\rho}\backslash\{e(t,\rho)\}\}$ are linearly independent.

Finally, after all updates, we have $CUT_{t,\rho} \subseteq In(t)$ for each $t \in T$ and each $\rho \in R_t(\delta_t)$, and $\mathrm{Rank}(\left(\tilde{f}_e^{\rho} : e \in CUT_{t,\rho}\right)) = w + \delta_t$. As the matrix $\left(\tilde{f}_e^{\rho} : e \in CUT_{t,\rho}\right)$ is a submatrix of $\tilde{F}_t^{\rho} \triangleq \left(\tilde{f}_e^{\rho} : e \in In(t)\right)$ with the same number of rows, it follows that $\mathrm{Rank}(\tilde{F}_t^{\rho}) = w + \delta_t$, i.e., $\Phi(t) \cap \Delta(t,\rho) = \{\underline{0}\}$.

For each error pattern $\eta \subseteq E$ satisfying $rank_t(\eta) < \delta_t$, there exists an error pattern $\rho \in R_t(\delta_t)$ such that $\eta \prec_t \rho$ from Proposition 1. This implies that $\Delta(t,\eta) \subseteq \Delta(t,\rho)$, and thus,

$$\Phi(t) \cap \Delta(t,\eta) \subseteq \Phi(t) \cap \Delta(t,\rho) = \{\underline{0}\}.$$

Now, we can say that $d_{\min}^{(t)} \geq \delta_t + 1$ for all $t \in T$, which, together with $d_{\min}^{(t)} \leq \delta_t + 1$ from Theorem 2, shows that $d_{\min}^{(t)} = \delta_t + 1$ for all $t \in T$.

The proof of 2): We just need to prove that if $|\mathcal{F}| \geq \sum_{t\in T} |R_t(\delta_t)|$, then

$$\Big|\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \backslash \cup_{t\in T} \cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]\Big| > 0.$$

Let $\dim(\tilde{\mathcal{L}}(In(i) \cup \{e'\})) = k$. For each $t \in T$ and $\rho \in R_t(\delta_t)$, if $e \in E_{t,\rho}$, then $e(t,\rho) \in In(i) \cup \{e'\}$, i.e., $\tilde{f}_{e(t,\rho)} \in \tilde{\mathcal{L}}(In(i) \cup \{e'\})$. Moreover, we know $\tilde{f}_{e(t,\rho)}^{\rho} \notin \tilde{\mathcal{L}}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\})$, i.e., $f_{e(t,\rho)}^{\rho} \notin \mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\})$, and $f_{e(t,\rho)}^{\rho} \notin \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})$. Together with $f_{e(t,\rho)}^{\rho^c} \in \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})$ and $\tilde{f}_{e(t,\rho)} = f_{e(t,\rho)}^{\rho} + f_{e(t,\rho)}^{\rho^c}$, this implies that

$$\tilde{f}_{e(t,\rho)} \notin \mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\}).$$

Therefore,

$$\dim\left(\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \cap [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]\right) \leq k - 1. \quad (5)$$

Consequently,

$$\begin{aligned}
&\Big|\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \backslash \cup_{t\in T} \cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]\Big| \\
=&\Big|\tilde{\mathcal{L}}(In(i) \cup \{e'\})\Big| - \Big|\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \cap \{\cup_{t\in T}\cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]\}\Big| && (6)\\
>&|\mathcal{F}|^k - \sum_{t\in T} \sum_{\rho\in R_t(\delta_t)} |\mathcal{F}|^{k-1} && (7)\\
\geq&|\mathcal{F}|^{k-1}[|\mathcal{F}| - \sum_{t\in T} |R_t(\delta_t)|\ ] \geq 0,
\end{aligned}$$

where the last step follows from $|\mathcal{F}| \geq \sum_{t\in T} |R_t(\delta_t)|$. For the inequality $(6) > (7)$, it is readily seen from (5) that $(6) \geq (7)$. It suffices to show $(6) > (7)$. It is not difficult to obtain that $(6) = (7)$, i.e.,

$$\Big|\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \cap \{\cup_{t\in T}\cup_{\substack{\rho\in R_t(\delta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i)\cup\{e'\})]\}\Big| = \sum_{t\in T} \sum_{\rho\in R_t(\delta_t)} |\mathcal{F}|^{k-1}$$

if and only if $|T| = 1,\ |R_t(\delta_t)| = 1$ and

$$\dim(\tilde{\mathcal{L}}(In(i) \cup \{e'\}) \cap [\mathcal{L}^{\rho}(CUT_{t,\rho}\backslash\{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]) = k - 1$$

with $e \in E_{t,\rho}$, where $R_t(\delta_t) = \{\rho\}$. However, it is impossible that $|R_t(\delta_t)| = 1$ because $\delta_t < C_t$. The proof is completed. ■

According to the known results, for the existence of the network error correction MDS codes, the size of the required base field is at least $\sum_{t\in T} \binom{|E|}{\delta_t}$. By Theorem 5, we can say that $\sum_{t\in T} |R_t(\delta_t)|$ is enough.

For any channel $e \in E$, if there exists a path from e to sink node t, then we call that e is connective with t.

Lemma 6: Let E_t be the set of channels which are connective with sink node $t \in T$. Then

$$\sum_{t\in T} |R_t(\delta_t)| \leq \sum_{t\in T} \binom{|E_t|}{\delta_t} \leq \sum_{t\in T} \binom{|E|}{\delta_t}.$$

Moreover, the necessary condition of the second inequality holding with equality is that there exists only one sink node in the network, i.e., $|T| = 1$.

Proof: Both inequalities are clear, and we will only consider the necessary condition of the second inequality holding with equality. Suppose that there are more than one sink node, and let t and t' be two distinct sink nodes. Obviously, there exists a channel e with $head(e) = t'$. That is, e is not connective with sink node t. This implies that $|E_t| < |E|$, and thus $\binom{|E_t|}{\delta_t} < \binom{|E|}{\delta_t}$, which shows that $\sum_{t\in T}\binom{|E_t|}{\delta_t} < \sum_{t\in T}\binom{|E|}{\delta_t}$. The lemma is proved. ■

From Theorem 5 and Lemma 6, we get the following corollary.

Corollary 7: If $|\mathcal{F}| \geq \sum_{t\in T}\binom{|E_t|}{\delta_t}$, then there exist linear network error correction MDS codes, i.e., for all $t \in T$,

$$d^{(t)}_{\min} = \delta_t + 1.$$

Example 1: Let G be a combination network [21, p.450] [20, p.26] with $N = 6$ and $k = 4$. That is, G is a single source multicast network, where there are $N = 6$ internal nodes, and one and only one channel from the source node s to each internal node. Arbitrary $k = 4$ internal nodes are connective with one and only one sink node, which implies that there are total $\binom{6}{4} = 15$ sink nodes. Thus, for G, we know that $|J| = 6$, $|T| = \binom{6}{4} = 15$, and $|E| = 6 + 4 \times \binom{6}{4} = 66$. It is evident that the minimum cut capacity C_t between s and any sink node t is 4. For example, Fig. 1 shows a combination network with $N = 3, k = 2$. Furthermore, let the information rate be $w = 2$, and thus $\delta_t = 2$ for each $t \in T$. Therefore, $|R_t(\delta_t)| = |R_t(2)| = 4 \times \binom{4}{2} = 24$ for each $t \in T$, and $\sum_{t\in T}|R_t(\delta_t)| = 15\times 24 = 360$. Nevertheless, $\sum_{t\in T}\binom{|E_t|}{\delta_t} = 15 \times \binom{8}{2} = 420$ and $\sum_{t\in T}\binom{|E|}{\delta_t} = 15 \times \binom{66}{2} = 32175$.

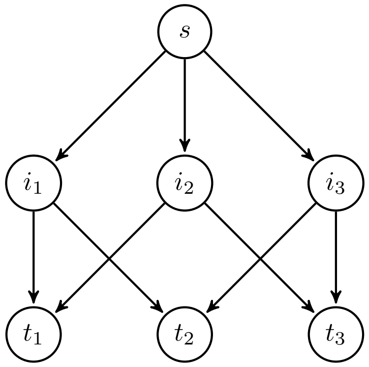

Fig. 1. Combination Network with $N = 3, k = 2$.

Now, we take into account the general network error correction codes, and give the following theorem.

Theorem 8: For any nonnegative integers β_t with $\beta_t \leq \delta_t$ for each $t \in T$, if $|\mathcal{F}| \geq \sum_{t\in T}|R_t(\beta_t)|$, then there exist linear network error correction codes satisfying for all $t \in T$,

$$d^{(t)}_{\min} \geq \beta_t + 1,$$

where $R_t(\beta_t)$ is the set of error patterns ρ satisfying $|\rho| = rank_t(\rho) = \beta_t$, that is,

$$R_t(\beta_t) = \{\text{error pattern } \rho : \ |\rho| = rank_t(\rho) = \beta_t\}.$$

The proof of this theorem is the same as that of Theorem 5 so long as replace δ_t by β_t, so the details are omitted.

The following conclusion shows that the required field size for constructing general linear network error correction codes is smaller than that for constructing network MDS codes.

Theorem 9: Let $\beta_t \leq \delta_t \leq \lfloor \frac{C_t}{2} \rfloor$, then $|R_t(\beta_t)| \leq |R_t(\delta_t)|$.

The proof of Theorem 9 is in Appendix A.

IV. The Constructive Algorithm of Linear Network Error Correction Codes

From the discussions in the last section, we propose the following Algorithm 1 for constructing a linear network error correction code with required error correction capability.

Algorithm 1 The algorithm for constructing a linear network error correction code with error correction capacity $d^{(t)}_{\min} \geq \beta_t$ for each $t \in T$.

Input: The single source multicast network $G = (V, E)$, the information rate $w \leq \min_{t\in T} C_t$, and the nonnegative integers $\beta_t \leq \delta_t$ for each $t \in T$.

Output: Extended global kernels (forming a linear network error correction code).

Initialization:

1) For each $t \in T$ and each $\rho \in R_t(\beta_t)$, find $(w + \beta_t)$ channel disjoint paths $\mathcal{P}_{t,\rho}$ from $In(s)$ or ρ' to t satisfying Corollary 4,
2) For each $t \in T$ and each $\rho \in R_t(\beta_t)$, initialize dynamic channel sets $CUT_{t,\rho} = In(s) \cup \rho' = \{d'_1, d'_2, \cdots, d'_w\} \cup \{e' : e \in \rho\}$, and the extended global encoding kernels $\tilde{f}_e = 1_e$ for all imaginary channels $e \in In(s) \cup E'$.

```
for each node $i \in V$ (according to the topological order of nodes) do
    for each channel $e \in Out(i)$ (according to an arbitrary order) do
        if $e \notin \cup_{t\in T} \cup_{\rho\in R_t(\beta_t)} E_{t,\rho}$ then
            $\tilde{f}_e = 1_e$,
            all $CUT_{t,\rho}$ remain unchanged.
        else if $e \in \cup_{t\in T} \cup_{\rho\in R_t(\beta_t)} E_{t,\rho}$ then
            choose $\tilde{g}_e \in \tilde{\mathcal{L}}(In(i) \cup \{e'\}) \backslash \cup_{t\in T} \cup_{\substack{\rho\in R_t(\beta_t):\\ e\in E_{t,\rho}}} [\mathcal{L}^{\rho}(CUT_{t,\rho} \backslash \{e(t,\rho)\}) + \mathcal{L}^{\rho^c}(In(i) \cup \{e'\})]$,
            if $\tilde{g}_e(e) = 0$ then
                $\tilde{f}_e = \tilde{g}_e + 1_e$,
            else
                $\tilde{f}_e = \tilde{g}_e(e)^{-1} \cdot \tilde{g}_e$.
            end if
            For those $CUT_{t,\rho}$ satisfying $e \in E_{t,\rho}$, update $CUT_{t,\rho} = \{CUT_{t,\rho} \backslash \{e(t,\rho)\}\} \cup \{e\}$; and for others, $CUT_{t,\rho}$ remain unchanged.
        end if
    end for
end for
```

Remark 2: Similar to the polynomial-time algorithm for constructing linear network codes in [5], our algorithm is a greedy one, too. The verification of Algorithm 1 is from the proof of Theorems 5 and 8. In particular, if we choose $\beta_t = \delta_t$ for all $t \in T$, then, by the proposed algorithm, we can construct a linear network error correction code that meets the refined Singleton bound with equality. That is, we can obtain a linear network error correction MDS code. On the other hand, if we choose $\beta_t = 0$ for each $t \in T$, then this algorithm degenerates into an algorithm for constructing linear network codes.

Next, we will analyze the time complexity of the proposed algorithm. First, from [5], we can determine $R_t(\beta_t)$ and find $(w+\beta_t)$ channel disjoint paths satisfying Lemma 3 in time $\mathcal{O}(\sum_{t\in T}\binom{|E|}{\beta_t}(w+\beta_t)|E|)$.

Both methods presented by Jaggi *et al.* [5] are used to analyze the time complexity of the main loop.

- If we use the method of Testing Linear Independent Quickly [5, III,A], the expected time complexity is at most

  $$\mathcal{O}\left(|E|\left[\sum_{t\in T}|R_t(\beta_t)|(w+\beta_t)(w+\frac{|E|+1}{2})\right]\right).$$

  After a simple calculation, the expected time complexity of the algorithm using the method of Testing Linear Independent Quickly is at most

  $$\mathcal{O}\Big(|E|(w+\beta_t) \cdot \left[\sum_{t\in T}\binom{|E|}{\beta_t}+\sum_{t\in T}|R_t(\beta_t)|(w+\frac{|E|+1}{2})\right]\Big).$$

- If we use the method of Deterministic Implementation [5, III,B], the time complexity of the main loop is at most

  $$\mathcal{O}\Big(|E|(w+\frac{|E|+1}{2}) \cdot \left[(\sum_{t\in T}|R_t(\beta_t)|)^2+\sum_{t\in T}|R_t(\beta_t)|(w+\beta_t)\right]\Big).$$

  Therefore, the total time complexity of the algorithm using the method of Deterministic Implementation is at most

  $$\mathcal{O}\Big(|E| \cdot\Big[(\sum_{t\in T}|R_t(\beta_t)|)^2(w+\frac{|E|+1}{2})+\sum_{t\in T}\binom{|E|}{\beta_t}(w+\beta_t)\Big]\Big).$$

As an example, we will apply Algorithm 1 to construct a network MDS code for a very simple network G_1 shown by Fig. 2.

Example 2: For the network G_1 shown by Fig. 2, let the topological order of all nodes be $s \prec i \prec t$, and the topological order of all channels be $e_1 \prec e_2 \prec e_3$. It is obvious that $C_t = 2$. Let $w = 1$, and thus $\delta_t = C_t - w = 1$. Furthermore, we have $R_t(\delta_t) = R_t(1) = \{\rho_1 = \{e_1\}, \rho_2 = \{e_2\}, \rho_3 = \{e_3\}\}$, and

$$\begin{aligned}
\mathcal{P}_{t,\rho_1} &= \{P_{t,\rho_1}^{(\delta_t)} = (e_1', e_1, e_3), P_{t,\rho_1}^{(w)} = (d', e_2)\},\\
E_{t,\rho_1} &= \{d', e_1', e_1, e_2, e_3\};\\
\mathcal{P}_{t,\rho_2} &= \{P_{t,\rho_2}^{(\delta_t)} = (e_2', e_2), P_{t,\rho_2}^{(w)} = (d', e_1, e_3)\},\\
E_{t,\rho_2} &= \{d', e_2', e_1, e_2, e_3\};\\
\mathcal{P}_{t,\rho_3} &= \{P_{t,\rho_3}^{(\delta_t)} = (e_3', e_3), P_{t,\rho_3}^{(w)} = (d', e_2)\},\\
E_{t,\rho_3} &= \{d', e_3', e_2, e_3\}.
\end{aligned}$$

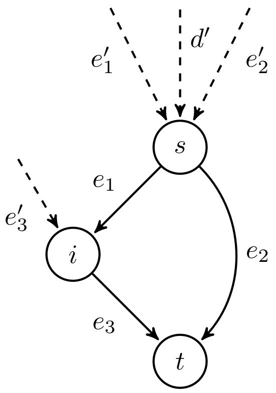

Fig. 2. Network G_1.

Let the base field be $\mathbb{Z}_3$. Initialize the dynamic channel sets $CUT_{t,\rho_1} = \{d', e_1'\}$, $CUT_{t,\rho_2} = \{d', e_2'\}$, $CUT_{t,\rho_3} = \{d', e_3'\}$, and

$$\tilde{f}_{d'} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \tilde{f}_{e_1'} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \tilde{f}_{e_2'} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \tilde{f}_{e_3'} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix},$$

which leads to $\dim(\tilde{\mathcal{L}}^{\rho_i}(CUT_{t,\rho_i})) = 2,\ (i = 1, 2, 3)$.

For the channel $e_1 \in Out(s)$, $e_1 \in E_{t,\rho_1} \cap E_{t,\rho_2}$ and

$$\begin{aligned}
&\tilde{\mathcal{L}}(\{d', e_1'\})\backslash[\mathcal{L}^{\rho_1}(\{d'\}) + \mathcal{L}^{\rho_1^c}(\{d', e_1'\})]\\
&\qquad\cup[\mathcal{L}^{\rho_2}(\{e_2'\}) + \mathcal{L}'^{\rho_2^c}(\{d', e_1'\})]\\
&= \left\langle\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right\rangle \Big\backslash \left\langle\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\right\rangle \cup \left\langle\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right\rangle + \left\langle\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right\rangle.
\end{aligned}$$

So we choose $\tilde{g}_{e_1} = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}$ because of

$$\begin{aligned}
\tilde{g}_{e_1} \in \tilde{\mathcal{L}}(\{d', e_1'\})&\backslash[\mathcal{L}^{\rho_1}(\{d'\}) + \mathcal{L}^{\rho_1^c}(\{d', e_1'\})]\\
&\cup[\mathcal{L}^{\rho_2}(\{e_2'\}) + \mathcal{L}'^{\rho_2^c}(\{d', e_1'\})].
\end{aligned}$$

And $\tilde{f}_{e_1} = \tilde{g}_{e_1}$, since $\tilde{g}_{e_1}(e_1) = 1$. Then update $CUT_{t,\rho_1} = \{d', e_1\}$, $CUT_{t,\rho_2} = \{e_1, e_2'\}$, and CUT_{t,ρ_3} remains unchanged.

For the channel $e_2 \in Out(s)$, $e_2 \in E_{t,\rho_1} \cap E_{t,\rho_2} \cap E_{t,\rho_3}$ and

$$\begin{aligned}
&\tilde{\mathcal{L}}(\{d', e_2'\})\backslash[\mathcal{L}^{\rho_1}(\{e_1\}) + \mathcal{L}^{\rho_1^c}(\{d', e_2'\})]\\
&\cup[\mathcal{L}^{\rho_2}(\{e_1\}) + \mathcal{L}^{\rho_2^c}(\{d', e_2'\})]\cup[\mathcal{L}^{\rho_3}(\{e_3'\}) + \mathcal{L}^{\rho_3^c}(\{d', e_2'\})]\\
&= \left\langle\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right\rangle \Big\backslash \left\langle\begin{pmatrix}1\\1\\0\\0\end{pmatrix}\right\rangle + \left\langle\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right\rangle\\
&\qquad\cup\left\langle\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\right\rangle \cup \left\langle\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\rangle + \left\langle\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right\rangle.
\end{aligned}$$

We choose $\tilde{g}_{e_2} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, since

$$\tilde{g}_{e_2} \in \tilde{\mathcal{L}}(\{d', e_2'\}) \backslash [\mathcal{L}^{\rho_1}(\{e_1\}) + \mathcal{L}^{\rho_1^c}(\{d', e_2'\})]$$
$$\cup [\mathcal{L}^{\rho_2}(\{e_1\}) + \mathcal{L}^{\rho_2^c}(\{d', e_2'\})] \cup [\mathcal{L}^{\rho_3}(\{e_3'\}) + \mathcal{L}^{\rho_3^c}(\{d', e_2'\})],$$

which, together with $\tilde{g}_{e_2}(e_2) = 1$, shows that $\tilde{f}_{e_2} = \tilde{g}_{e_2}$. Then, update $CUT_{t,\rho_1} = \{e_2, e_1\}$, $CUT_{t,\rho_2} = \{e_1, e_2\}$, and $CUT_{t,\rho_3} = \{e_2, e_3'\}$.

For the channel $e_3 \in Out(i)$, $e_3 \in E_{t,\rho_1} \cap E_{t,\rho_2} \cap E_{t,\rho_3}$ and

$$\tilde{\mathcal{L}}(\{e_1, e_3'\}) \backslash [\mathcal{L}^{\rho_1}(\{e_2\}) + \mathcal{L}^{\rho_1^c}(\{e_1, e_3'\})]$$
$$\cup [\mathcal{L}^{\rho_2}(\{e_2\}) + \mathcal{L}^{\rho_2^c}(\{e_1, e_3'\})] \cup [\mathcal{L}^{\rho_3}(\{e_2\}) + \mathcal{L}^{\rho_3^c}(\{e_1, e_3'\})]$$
$$= \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \Big\backslash \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$
$$\cup \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \cup \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle.$$

We select $\tilde{g}_{e_3} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ satisfying

$$\tilde{g}_{e_3} \in \tilde{\mathcal{L}}(\{e_1, e_3'\}) \backslash [\mathcal{L}^{\rho_1}(\{e_2\}) + \mathcal{L}^{\rho_1^c}(\{e_1, e_3'\})]$$
$$\cup [\mathcal{L}^{\rho_2}(\{e_2\}) + \mathcal{L}^{\rho_2^c}(\{e_1, e_3'\})] \cup [\mathcal{L}^{\rho_3}(\{e_2\}) + \mathcal{L}^{\rho_3^c}(\{e_1, e_3'\})].$$

It follows that $\tilde{f}_{e_3} = \tilde{g}_{e_3}$ from $\tilde{g}_{e_3}(e_3) = 1$, and update $CUT_{t,\rho_1} = CUT_{t,\rho_2} = CUT_{t,\rho_3} = \{e_2, e_3\} \subseteq In(t)$.

The decoding matrix at t is $\tilde{F}_t = (\tilde{f}_{e_2} \ \tilde{f}_{e_3}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$. It is easy to check that $\Phi(t) \cap \Delta(t, \rho_i) = \{\underline{0}\}$ for $i = 1, 2, 3$. Further, let $\rho = \{e_1, e_2\}$. Then $rank_t(\rho) = 2$ and $\Phi(t) \cap \Delta(t, \rho) \neq \{\underline{0}\}$, which means $d_{\min}^{(t)} = 2 = \delta_t + 1$. That is, $\{\tilde{f}_{e_1}, \tilde{f}_{e_2}, \tilde{f}_{e_3}\}$ forms a global description of a linear network error correction MDS code for the network G_1.

V. Random Linear Network Error Correction Coding

Random network coding was originally proposed in [6]. When a node (maybe the source node s) receives the messages from its all incoming channels, for each outgoing channel, it selects the encoding coefficients uniformly at random over the base field $\mathcal{F}$, uses them to encode the messages and transmits the encoded messages over the outgoing channel. In other words, the local encoding coefficients $k_{d,e}$ are independently, uniformly distributed random variables on the base field $\mathcal{F}$. The performance analysis of random linear network coding is very important in theory and applications. In this section, we will investigate the error correction capability of random linear network coding. We first consider random linear network error correction MDS codes. Before the discussion, we give the following definitions.

Definition 8: Let G be a single source multicast network, $\mathbf{C}$ be a random linear network error correction code on G, and $d_{\min}^{(t)}$ be the minimum distance at sink node t of $\mathbf{C}$.

- $P_{ec}(t) \triangleq Pr(\{\dim(\Phi(t)) < w\} \cup \{d_{\min}^{(t)} < \delta_t + 1\})$ is called the failure probability of random linear network error correction MDS coding for sink node t.
- $P_{ec} \triangleq Pr(\{\mathbf{C}$ is not regular$\} \cup \{\exists\, t \in T$such that $d_{\min}^{(t)} < \delta_t + 1\})$ is called the failure probability of random linear network error correction MDS coding for network G, that is the probability that network MDS codes are not constructed by the random method.

In order to evaluate these two failure probabilities, the following lemma is useful.

Lemma 10 ([8, Lemma 1], [22]): Let $\mathcal{L}$ be an n dimensional linear space over a finite field $\mathcal{F}$, $\mathcal{L}_0$, $\mathcal{L}_1$ be two subspaces of $\mathcal{L}$ of dimensions k_0, k_1, respectively, and $\langle \mathcal{L}_0 \cup \mathcal{L}_1 \rangle = \mathcal{L}$. Let $l_1, \ l_2, \ \cdots, \ l_m$ $(m = n - k_0)$ be m independently and uniformly distributed random vectors taking values in $\mathcal{L}_1$. Then

$$Pr(\dim(\langle \mathcal{L}_0 \cup \{l_1, \ l_2, \ \cdots, \ l_m\} \rangle) = n) = \prod_{i=1}^{m} \left(1 - \frac{1}{|\mathcal{F}|^i}\right).$$

Theorem 11: Let G be a single source multicast network, and $w \leq \min_{t \in T} C_t$. Using random method to construct a linear network error correction MDS code, then

- for each $t \in T$, the failure probability of random linear network error correction MDS coding for t satisfies
$$P_{ec}(t) < 1 - \left(1 - \frac{|R_t(\delta_t)|}{|\mathcal{F}| - 1}\right)^{|J|+1};$$
- the failure probability of random linear network error correction MDS coding for the network G satisfies
$$P_{ec} < 1 - \left(1 - \frac{\sum_{t \in T} |R_t(\delta_t)|}{|\mathcal{F}| - 1}\right)^{|J|+1},$$
where J is the set of the internal nodes in G.

Proof: For the single source multicast network $G = (V, E)$, s is the single source node, T is the set of the sink nodes, $J = V - \{s\} - T$ is the set of the internal nodes, and E is the set of all channels. Let $\tilde{G} = (\tilde{V}, \tilde{E})$ be the extended network of G.

For each sink node $t \in T$ and each error pattern $\rho \in R_t(\delta_t)$, Corollary 4 implies that there are $(w + \delta_t)$ channel disjoint paths from either $In(s)$ or ρ' to t satisfying the properties that (1) there exist exactly δ_t channel disjoint paths from ρ' to t, and w channel disjoint paths from $In(s)$ to t; (2) each of these δ_t paths from ρ' to t starts with a channel $e' \in \rho'$ and passes through the corresponding channel $e \in \rho$. Denote by $\mathcal{P}_{t,\rho}$ the set of $(w + \delta_t)$ channel disjoint paths satisfying these properties and $E_{t,\rho}$ denotes the set of all channels in $\mathcal{P}_{t,\rho}$.

Note that the event "$\{\dim(\Phi(t)) = w\} \cap \{d_{\min}^{(t)} = \delta_t + 1\}$" is equivalent to the event "$\{\dim(\Phi(t)) = w\} \cap \{\forall\ \rho \in R_t(\delta_t) : \Phi(t) \cap \Delta(t, \rho) = \{\underline{0}\}\}$", and furthermore, the event "$\forall\ \rho \in R_t(\delta_t) : \mathrm{Rank}(\tilde{F}_t^{\rho}) = w + \delta_t$" implies the event "$\{\dim(\Phi(t)) = w\} \cap \{\forall\ \rho \in R_t(\delta_t) : \Phi(t) \cap \Delta(t, \rho) = \{\underline{0}\}\}$". Thus, we consider the following probability:

$$Pr(\cap_{\rho \in R_t(\delta_t)} \mathrm{Rank}(\tilde{F}_t^{\rho}) = w + \delta_t).$$

For the network G, let an ancestral order of nodes be

$$s \prec i_1 \prec i_2 \prec \cdots \prec i_{|J|} \prec T.$$

During our discussion, we use the concept of cuts of the paths similar to the dynamic set $CUT_{t,\rho}$ as mentioned above. The first cut is $CUT_{t,\rho,0} = In(s) \cup \{e' : e \in \rho\}$, i.e., the

w imaginary message channels $d'_1, d'_2, \cdots, d'_w$ and imaginary error channels corresponding to the channels in ρ. At node s, the next $CUT_{t,\rho,1}$ is formed from $CUT_{t,\rho,0}$ by replacing those channels in $\{In(s) \cup \{e' : e \in Out(s)\}\} \cap CUT_{t,\rho,0}$ by their respective next channels in the paths. These new channels are in $Out(s) \cap E_{t,\rho}$. Other channels remain the same as in $CUT_{t,\rho,0}$. At node i_1, the next cut $CUT_{t,\rho,2}$ is formed from $CUT_{t,\rho,1}$ by replacing those channels in $\{In(i_1) \cup \{e' : e \in Out(i_1)\}\} \cap CUT_{t,\rho,1}$ by their respective next channels in the paths. These new channels are in $Out(i_1) \cap E_{t,\rho}$. Other channels remain the same as in $CUT_{t,\rho,1}$. Subsequently, once $CUT_{t,\rho,k}$ is defined, $CUT_{t,\rho,k+1}$ is formed from $CUT_{t,\rho,k}$ by the same method. By induction, all cuts $CUT_{t,\rho,k}$ for $t \in T$, $\rho \in R_t(\delta_t)$, and $k = 0, 1, 2, \cdots, |J| + 1$ can be defined. Moreover, for each $CUT_{t,\rho,k}$, we divide $CUT_{t,\rho,k}$ into two disjoint parts $CUT^{in}_{t,\rho,k}$ and $CUT^{out}_{t,\rho,k}$ as follows:

$$CUT^{in}_{t,\rho,k} = \{e : e \in CUT_{t,\rho,k} \cap In(i_k)\},$$
$$CUT^{out}_{t,\rho,k} = \{e : e \in CUT_{t,\rho,k} \backslash CUT^{in}_{t,\rho,k}\}.$$

Define $(w + \delta_t) \times (w + \delta_t)$ matrix $\tilde{F}^{\rho(k)}_t = (\tilde{f}^{\rho}_e : e \in CUT_{t,\rho,k})$ for $k = 0, 1, \cdots, |J|+1$. If $\mathrm{Rank}(\tilde{F}^{\rho(k)}_t) < w+\delta_t$, we call that we have a failure at $CUT_{t,\rho,k}$. Let $\Gamma^{(t,\rho)}_k$ represent the event "$\mathrm{Rank}(\tilde{F}^{\rho(k)}_t) = w+\delta_t$". Furthermore, let $|J| = m$, and note that $\tilde{F}^{\rho(m+1)}_t$ is a submatrix of $\tilde{F}^{\rho}_t$. It follows that the event "$\forall \rho \in R_t(\delta_t), \mathrm{Rank}(\tilde{F}^{\rho(m+1)}_t) = w + \delta_t$" implies the event "$\forall \rho \in R_t(\delta_t), \mathrm{Rank}(\tilde{F}^{\rho}_t) = w + \delta_t$". Therefore,

$$\begin{aligned}
&1 - P_{ec}(t)\\
=&Pr(\{\dim(\Phi(t)) = w\} \cap \{d^{(t)}_{\min} = \delta_t + 1\})\\
=&Pr(\{\dim(\Phi(t)) = w\} \cap \{\cap_{\rho \in R_t(\delta_t)} \Phi(t) \cap \Delta(t,\rho) = \{\underline{0}\}\})\\
\geq&Pr(\cap_{\rho \in R_t(\delta_t)} \mathrm{Rank}(\tilde{F}^{\rho}_t) = w + \delta_t)\\
\geq&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1}).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1})\\
\geq&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1}, \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m}, \cdots, \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0})\\
\geq&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m}) \cdots\\
&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0}) Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0})\\
=&\prod_{k=0}^{m} Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k}), \qquad (8)
\end{aligned}$$

where (8) follows from

$$\begin{aligned}
&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0})\\
=&Pr(\cap_{\rho \in R_t(\delta_t)} \mathrm{Rank}((\tilde{f}^{\rho}_e : e \in In(s) \cup \rho')) = w + \delta_t)\\
=&Pr(\mathrm{Rank}(I_{w+\delta_t}) = w + \delta_t) \equiv 1.
\end{aligned}$$

For each channel $e \in E$, let $e \in Out(i_k)$. Let $\tilde{g}_e$ be an independently and uniformly distributed random vector taking values in $\tilde{\mathcal{L}}(In(i_k))$. In other words, if $In(i_k) = \{d_1, d_2, \cdots, d_l\}$, then

$$\tilde{g}_e = k_{d_1,e}\tilde{f}_{d_1} + k_{d_2,e}\tilde{f}_{d_2} + \cdots + k_{d_l,e}\tilde{f}_{d_l},$$

where $k_{d_j,e}$ $(j = 1, 2, \cdots, l)$ are independently and uniformly distributed random variables taking values in the base field $\mathcal{F}$. It follows that $\tilde{g}^{\rho}_e = k_{d_1,e}\tilde{f}^{\rho}_{d_1} + k_{d_2,e}\tilde{f}^{\rho}_{d_2} + \cdots + k_{d_l,e}\tilde{f}^{\rho}_{d_l}$ is also an independently and uniformly distributed random vector taking values in $\tilde{\mathcal{L}}^{\rho}(In(i_k))$. We always define $\tilde{f}_e = \tilde{g}_e + 1_e$. Therefore, for all $e \in E_{t,\rho} \cap Out(i_k)$ with $e(t,\rho) \in CUT^{in}_{t,\rho,k}$, i.e., $e \notin \rho$, it is shown that $\tilde{f}^{\rho}_e = \tilde{g}^{\rho}_e$ because of $e \notin \rho$. Thus, $\tilde{f}^{\rho}_e$ is an independently and uniformly distributed random vector taking values in $\tilde{\mathcal{L}}^{\rho}(In(i_k))$. Otherwise $e \in E_{t,\rho} \cap Out(i_k)$ with $e(t,\rho) \in CUT^{out}_{t,\rho,k}$, that is, $e(t,\rho) = e'$, then, $\tilde{f}^{\rho}_e$ and $\{\tilde{f}^{\rho}_d : d \in CUT_{t,\rho,k} \backslash e(t,\rho)\}$ are always linearly independent, since $\tilde{f}^{\rho}_e(e) = 1$ and $\tilde{f}^{\rho}_d(e) = 0$ for all $d \in CUT_{t,\rho,k} \backslash e(t,\rho)$.

Applying Lemma 10, we derive

$$\begin{aligned}
&Pr(\Gamma^{(t,\rho)}_{k+1} | \Gamma^{(t,\rho)}_{k}) = \prod_{i=1}^{|CUT^{in}_{t,\rho,k}|} \left(1 - \frac{1}{|\mathcal{F}|^i}\right)\\
\geq&\prod_{i=1}^{w+\delta_t} \left(1 - \frac{1}{|\mathcal{F}|^i}\right) > 1 - \sum_{i=1}^{w+\delta_t} \frac{1}{|\mathcal{F}|^i}\\
>&1 - \sum_{i=1}^{\infty} \frac{1}{|\mathcal{F}|^i} = 1 - \frac{1}{|\mathcal{F}| - 1}.
\end{aligned}$$

Consequently, for each k $(0 \leq k \leq m)$, one has

$$\begin{aligned}
&Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k})\\
=&1 - Pr(\cup_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)^c}_{k+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k})\\
\geq&1 - \sum_{\rho \in R_t(\delta_t)} Pr(\Gamma^{(t,\rho)^c}_{k+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k})\\
=&1 - \sum_{\rho \in R_t(\delta_t)} Pr(\Gamma^{(t,\rho)^c}_{k+1} | \Gamma^{(t,\rho)}_{k})\\
>&1 - \sum_{\rho \in R_t(\delta_t)} \frac{1}{|\mathcal{F}| - 1}\\
=&1 - \frac{|R_t(\delta_t)|}{|\mathcal{F}| - 1}.
\end{aligned}$$

Combining the above inequalities, we have

$$\begin{aligned}
1 - P_{ec}(t) \geq& \prod_{k=0}^{m} Pr(\cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k+1} | \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k})\\
>& \left(1 - \frac{|R_t(\delta_t)|}{|\mathcal{F}| - 1}\right)^{m+1}.
\end{aligned}$$

That is,

$$P_{ec}(t) < 1 - \left(1 - \frac{|R_t(\delta_t)|}{|\mathcal{F}| - 1}\right)^{m+1}.$$

Next,

$$\begin{aligned}
&1 - P_{ec} \geq Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} Rank(\tilde{F}^{\rho}_t) = w + \delta_t)\\
\geq&Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1}, \cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m}, \cdots,\\
&\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0})\\
\geq&Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m+1} | \cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m})\\
&\cdot Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m} | \cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{m-1}) \cdots\\
&\cdot Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{1} | \cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{0}) \qquad (9)\\
=&\prod_{k=0}^{m} Pr(\cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k+1} | \cap_{t \in T} \cap_{\rho \in R_t(\delta_t)} \Gamma^{(t,\rho)}_{k}),
\end{aligned}$$

where (9) follows from $Pr(\cap_{t\in T}\cap_{\rho\in R_t(\delta_t)}\Gamma_0^{(t,\rho)})\equiv 1$.

Furthermore, for each k $(0\le k\le m)$,

$$\begin{aligned}
&Pr(\cap_{t\in T}\cap_{\rho\in R_t(\delta_t)}\Gamma_{k+1}^{(t,\rho)}|\cap_{t\in T}\cap_{\rho\in R_t(\delta_t)}\Gamma_k^{(t,\rho)})\\
=&1-Pr(\cup_{t\in T}\cup_{\rho\in R_t(\delta_t)}\Gamma_{k+1}^{(t,\rho)^c}|\cap_{t\in T}\cap_{\rho\in R_t(\delta_t)}\Gamma_k^{(t,\rho)})\\
\ge&1-\sum_{t\in T}\sum_{\rho\in R_t(\delta_t)}Pr(\Gamma_{k+1}^{(t,\rho)^c}|\Gamma_k^{(t,\rho)})\\
=&1-\sum_{t\in T}\sum_{\rho\in R_t(\delta_t)}[1-Pr(\Gamma_{k+1}^{(t,\rho)}|\Gamma_k^{(t,\rho)})]\\
>&1-\sum_{t\in T}\sum_{\rho\in R_t(\delta_t)}\frac{1}{|\mathcal{F}|-1}\\
=&1-\frac{\sum_{t\in T}|R_t(\delta_t)|}{|\mathcal{F}|-1}. \qquad (10)
\end{aligned}$$

Combining the inequalities (9) and (10), we have

$$1-P_{ec}>\left(1-\frac{\sum_{t\in T}|R_t(\delta_t)|}{|\mathcal{F}|-1}\right)^{m+1},$$

that is,

$$P_{ec}<1-\left(1-\frac{\sum_{t\in T}|R_t(\delta_t)|}{|\mathcal{F}|-1}\right)^{m+1}.$$

The proof is completed. ■

Applying Lemma 6 to Theorem 11, we derive the following corollary.

Corollary 12: The failure probability $P_{ec}(t)$ of random linear network error correction MDS coding for each $t\in T$ satisfies

$$P_{ec}(t)<1-\left(1-\frac{\binom{|E_t|}{\delta_t}}{|\mathcal{F}|-1}\right)^{|J|+1}\le 1-\left(1-\frac{\binom{|E|}{\delta_t}}{|\mathcal{F}|-1}\right)^{|J|+1}.$$

The failure probability P_{ec} of random linear network error correction MDS coding for the network G satisfies

$$\begin{aligned}
P_{ec}<&1-\left(1-\frac{\sum_{t\in T}\binom{|E_t|}{\delta_t}}{|\mathcal{F}|-1}\right)^{|J|+1}\\
\le&1-\left(1-\frac{\sum_{t\in T}\binom{|E|}{\delta_t}}{|\mathcal{F}|-1}\right)^{|J|+1}.
\end{aligned}$$

However, in practice, we sometimes need general linear network error correction codes instead of the network MDS codes. That is, we only need the codes satisfying that its minimum distance $d_{\min}^{(t)}\ge\beta_t$, where $\beta_t\le\delta_t$ is a nonnegative integer. The part of reason is that usually the field size required by general linear network error correction codes is smaller than that of network MDS codes. Hence, we should also discuss the random method for the general linear network error correction codes. Similarly, we define the failure probabilities for random linear network error correction codes as follows.

Definition 9: Let G be a single source multicast network, $\mathbf{C}$ be a random linear network error correction code on G, and $d_{\min}^{(t)}$ be the minimum distance at sink node t. Define that

- $P_{ec}(t,\beta_t)\triangleq Pr(\{\dim(\Phi(t))<w\}\cup\{d_{\min}^{(t)}<\beta_t+1\})$, that is the probability that the code $\mathbf{C}$ cannot either be decoded or satisfy that the error correction capacity $d_{\min}^{(t)}\ge\beta_t+1$ at the sink node t;
- $P_{ec}(\beta_t)\triangleq Pr(\{\ \mathbf{C}$ is not regular $\}\cup\{\exists\, t\in T$ such that $d_{\min}^{(t)}<\beta_t+1\})$, that is the probability that the regular linear network error correction codes with $d_{\min}^{(t)}\ge\beta_t+1$ cannot be constructed by the random method.

Using the similar method to prove Theorem 11, and combining it with the method to prove the random linear network coding with proper redundancy [7, Theorem 2], we can get the following results.

Theorem 13: Let G be a single source multicast network, the minimum cut capacity for sink node $t\in T$ be C_t and the information rate be w symbols per unit time satisfying $w\le\min_{t\in T}C_t$. Using random method to construct a linear network error correction code, then

- for each $t\in T$ and $\beta_t\le\delta_t$,
$$P_{ec}(t,\beta_t)\le\frac{|R_t(\beta_t)|\binom{\delta_t-\beta_t+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{\delta_t-\beta_t+1}};$$
- for the network G,
$$P_{ec}(\beta_t)\le\sum_{t\in T}\frac{|R_t(\beta_t)|\binom{\delta_t-\beta_t+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{\delta_t-\beta_t+1}}.$$

Remark 3: Both Theorems 11 and 13 above imply that these failure probabilities can become arbitrarily small when the size of the base field $\mathcal{F}$ is sufficiently large.

Balli, Yan, and Zhang [7] used $D_{\min}^{(t)}$ to denote the minimum distance of random linear network error correction code at a sink node $t\in T$. Obviously, the refined Singleton bound tells us that $D_{\min}^{(t)}$ takes values in $\{0,1,2,\cdots,\delta_t+1\}$. Furthermore, they studied the probability mass function of $D_{\min}^{(t)}$. For a code with the minimum distance $d_{\min}^{(t)}$ at sink node t, $\delta_t+1-d_{\min}^{(t)}$ is called the degradation of the code at t. Then they presented the following conclusions.

Proposition 3 ([7, Theorem 4]): For single source multicast over an acyclic network G, let the minimum cut capacity for sink node $t\in T$ be C_t, let the information rate be w symbols per unit time, let $\delta_t=C_t-w$ be the redundancy of the code for the sink node $t\in T$. For a given $d\ge 0$, the linear random network code satisfies:

$$Pr(D_{\min}^{(t)}<\delta_t+1-d)\le\frac{\binom{|E|}{\delta_t-d}\binom{d+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{d+1}}.$$

Furthermore, the probability that random linear network code has minimum distance at least δ_t+1-d at all sinks $t\in T$ is lower bounded by,

$$Pr(D_{\min}^{(t)}\ge\delta_t+1-d,\forall\ t\in T)\ge 1-\sum_{t\in T}\frac{\binom{|E|}{\delta_t-d}\binom{d+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{d+1}}.$$

This proposition can lead to an upper bound on the field size required for the existence of linear network error correction codes with degradation at most d.

Proposition 4 ([7, Corollary 1]): If the field size satisfies the following condition:

$$|\mathcal{F}|\ge 2+\left(\sum_{t\in T}\binom{|E|}{\delta_t-d}\binom{d+|J|+1}{|J|}\right)^{\frac{1}{d+1}},$$

then there exists a code having degradation at most d at all sinks $t \in T$.

In the same way, applying Theorem 13, we can also get a probability mass function of $D_{\min}^{(t)}$.

Corollary 14: For a single source multicast network $G = (V, E)$, let the minimum cut capacity for sink node $t \in T$ be C_t, the information rate be w symbols per unit time satisfying $w \le \min_{t \in T} C_t$, and $\delta_t = C_t - w$ be the redundancy of the code for sink $t \in T$. For a given $d \ge 0$, the random linear network error correction codes satisfy:

$$Pr(D_{\min}^{(t)} < \delta_t + 1 - d) \le \frac{|R_t(\delta_t - d)|\binom{d+|J|+1}{|J|}}{(|\mathcal{F}| - 1)^{d+1}},$$

and

$$\begin{aligned} &Pr(D_{\min}^{(t)} \ge \delta_t + 1 - d, \forall\ t \in T) \\ \ge &1 - \sum_{t \in T} \frac{|R_t(\delta_t - d)|\binom{d+|J|+1}{|J|}}{(|\mathcal{F}| - 1)^{d+1}}. \end{aligned}$$

This corollary also leads to an upper bound on the field size required for the existence of linear network error correction codes with degradation at most d. On the other hand, Theorem 8 shows that the required field size satisfies $|\mathcal{F}| \ge \sum_{t \in T} |R(\beta_t)|$. Therefore, we derive the following result.

Corollary 15: If the size of the base field $\mathcal{F}$ satisfies the following condition:

$$\begin{aligned} |\mathcal{F}| \ge \min\Big\{ &\sum_{t \in T} |R_t(\delta_t - d)|, \\ &2 + \left[\sum_{t \in T} |R(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}} \Big\}, \end{aligned}$$

then there exists a regular linear network error correction code having degradation at most d at all sink nodes $t \in T$.

When $d = 0$, it is readily seen that

$$\begin{aligned} &\sum_{t \in T} |R_t(\delta_t - d)| \\ = &\sum_{t \in T} |R_t(\delta_t)| \\ < &2 + (|J| + 1) \sum_{t \in T} |R_t(\delta_t)| \\ = &2 + \left[\sum_{t \in T} |R_t(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}}. \end{aligned}$$

This means that, for network MDS codes, Corollary 15 cannot give a smaller field size required. But, for $d \ge 1$, the size bounds $2 + \left[\sum_{t \in T} |R_t(\delta_t - d)|\binom{d+|J|+1}{|J|}\right]^{\frac{1}{d+1}}$ and $\sum_{t \in T} |R_t(\delta_t - d)|$ have no deterministic relations. We will illustrate this point through the following example.

Example 3: For network G_2 shown by Fig. 3 below, let $w = 2$. Then $\delta_t = C_t - w = 2$.

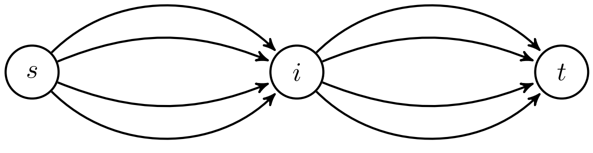

Fig. 3. Network G_2 with $|T| = 1$, $|J| = 1$, $C_t = 4$.

- In the case $d = 0$, it is clear that

$$\begin{aligned} &2 + \left[\sum_{t \in T} |R_t(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}} \\ = &2 + 2|R_t(2)| > |R_t(2)|. \end{aligned}$$

- In the case $d = 1$, a simple calculation gives

$$\sum_{t \in T} |R_t(\delta_t - d)| = |R_t(1)| = 8,$$

  and

$$\begin{aligned} &2 + \left[\sum_{t \in T} |R_t(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}} \\ = &2 + \sqrt{24} < 2 + 5 = 7. \end{aligned}$$

  This shows that in this case

$$\begin{aligned} &2 + \left[\sum_{t \in T} |R_t(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}} \\ < &\sum_{t \in T} |R_t(\delta_t - d)|. \end{aligned}$$

Nevertheless, for the network G_3 shown by Fig. 4, let $w = 2$, which shows $\delta_t = C_t - w = 2$.

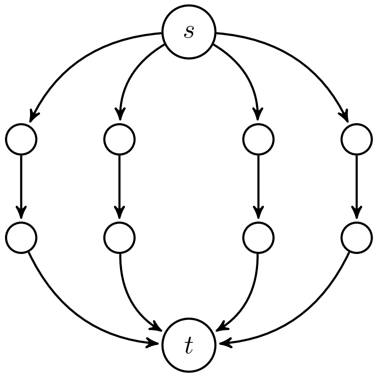

Fig. 4. Network G_3 with $|T| = 1$, $|J| = 8$, $C_t = 4$.

- In the case $d = 0$, obviously,

$$\begin{aligned} &2 + \left[\sum_{t \in T} |R_t(\delta_t - d)| \binom{d + |J| + 1}{|J|}\right]^{\frac{1}{d+1}} \\ = &2 + 9|R_t(2)| > |R_t(2)|. \end{aligned}$$

- In the case $d = 1$, after a simple calculation, we deduce that

$$\sum_{t \in T} |R_t(\delta_t - d)| = |R_t(1)| = 12,$$

and

$$
\begin{aligned}
&2+\left[\sum_{t\in T}|R_t(\delta_t-d)|\binom{d+|J|+1}{|J|}\right]^{\frac{1}{d+1}}\\
=&2+\left[|R_t(1)|\binom{1+8+1}{8}\right]^{\frac{1}{2}}=2+(12\times 45)^{\frac{1}{2}}\geq 20.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&2+\left[\sum_{t\in T}|R_t(\delta_t-d)|\binom{d+|J|+1}{|J|}\right]^{\frac{1}{d+1}}\\
>&\sum_{t\in T}|R_t(\delta_t-d)|.
\end{aligned}
$$

VI. Conclusions

In this paper, using the extended global encoding kernels proposed by Zhang in [12], we can prove the refined Singleton bound in network error correction coding more easily, and give a constructive proof to show that this bound is tight, that is, we construct network MDS codes which meet this bound with equality. As a consequence of this proof, an algorithm is designed to construct linear network error correction codes, especially network MDS codes. The time complexity of the proposed algorithm is analyzed. It is shown that the required field size for the existence of linear network error correction codes can become smaller than the previously known results, and even much smaller in some cases.

For random linear network error correction coding, the upper bounds on the failure probabilities for network MDS codes and general linear network error correction codes are obtained. And we slightly improve on the probability mass function of the minimum distance of the random linear network error correction codes introduced in [7], as well as the upper bound on the field size required for the existence of linear network error correction codes with degradation at most d.

Appendix A
Proof of Theorem 9

Proof: We choose an error pattern $\rho_1 \in R_t(\beta_t)$ arbitrarily, that is, the chosen error pattern ρ_1 satisfies $|\rho_1| = rank_t(\rho_1) = \beta_t$. Then we can extend ρ_1 to an error pattern ρ_1' with $\rho_1 \subseteq \rho_1'$ and $|\rho_1'| = rank_t(\rho_1') = C_t$, since the minimum cut capacity between s and t is C_t. Define two sets as follows:

$$\Omega_{1,\beta_t} = \{\text{error pattern } \rho \subseteq \rho_1' : \rho \in R_t(\beta_t)\}$$

and

$$\Omega_{1,\delta_t} = \{\text{error pattern } \rho' \subseteq \rho_1' : \rho' \in R_t(\delta_t)\}.$$

From the above definitions, we have

$$|\Omega_{1,\beta_t}| = \binom{C_t}{\beta_t} \text{ and } |\Omega_{1,\delta_t}| = \binom{C_t}{\delta_t}.$$

Note that $\beta_t \leq \delta_t \leq \lfloor\frac{C_t}{2}\rfloor$ implies $\binom{C_t}{\beta_t} \leq \binom{C_t}{\delta_t}$. In other words, for each $\rho \in \Omega_{1,\beta_t}$, there exists an error pattern $\rho' \in \Omega_{1,\delta_t}$ such that ρ is covered by ρ', i.e., $\rho \subseteq \rho'$, and $\theta' \neq \eta'$ for any distinct $\theta, \eta \in \Omega_{1,\beta_t}$.

Again, choose an error pattern $\rho_2 \in R_t(\beta_t)\backslash\Omega_{1,\beta_t}$ arbitrarily. In the same way as for ρ_1, ρ_2 can be extended to an error pattern ρ_2' with $\rho_2 \subseteq \rho_2'$ and $|\rho_2'| = rank_t(\rho_2') = C_t$. Define the next two sets:

$$\Omega_{2,\beta_t} = \{\text{error pattern } \rho \subseteq \rho_2' : \rho \in R_t(\beta_t), \rho \nsubseteq \rho_1' \cap \rho_2'\},$$

and

$$\Omega_{2,\delta_t} = \{\text{error pattern } \rho' \subseteq \rho_2' : \rho' \in R_t(\delta_t), \rho' \nsubseteq \rho_1' \cap \rho_2'\}.$$

Obviously, for all $\rho \in \Omega_{2,\beta_t}$ and $\rho' \in \Omega_{2,\delta_t}$, we have $\rho \notin \Omega_{1,\beta_t}$ and $\rho' \notin \Omega_{1,\delta_t}$. This means that $\Omega_{1,\beta_t} \cap \Omega_{2,\beta_t} = \emptyset$ and $\Omega_{1,\delta_t} \cap \Omega_{2,\delta_t} = \emptyset$. Let $|\rho_1' \cap \rho_2'| = k_{1,2}$. Then

$$|\Omega_{2,\beta_t}| = \binom{C_t}{\beta_t} - \binom{k_{1,2}}{\beta_t} \text{ and } |\Omega_{2,\delta_t}| = \binom{C_t}{\delta_t} - \binom{k_{1,2}}{\delta_t}.$$

We adopt the convention that $\binom{a}{b} = 0$ for $a < b$.

Similarly, we choose an error pattern $\rho_3 \in R_t(\beta_t)\backslash\Omega_{1,\beta_t} \cup \Omega_{2,\beta_t}$, and extend ρ_3 to an error pattern ρ_3' with $\rho_3 \subseteq \rho_3'$ and $|\rho_3'| = rank_t(\rho_3') = C_t$. Define

$$\Omega_{3,\beta_t} = \{\rho \subseteq \rho_3' : \rho \in R_t(\beta_t), \rho \nsubseteq \{\rho_1' \cup \rho_2'\} \cap \rho_3'\},$$

and

$$\Omega_{3,\delta_t} = \{\rho' \subseteq \rho_3' : \rho' \in R_t(\delta_t), \rho' \nsubseteq \{\rho_1' \cup \rho_2'\} \cap \rho_3'\}.$$

We claim that for all $\rho \in \Omega_{3,\beta_t}$ and $\rho' \in \Omega_{3,\delta_t}$, $\rho \notin \Omega_{1,\beta_t} \cup \Omega_{2,\beta_t}$ and $\rho' \notin \Omega_{1,\delta_t} \cup \Omega_{2,\delta_t}$. Conversely, suppose that $\rho \in \cup_{i=1}^{2}\Omega_{i,\beta_t}$ (resp. $\rho' \in \cup_{i=1}^{2}\Omega_{i,\delta_t}$). Together with $\rho \in \Omega_{3,\beta_t}$ (resp. $\rho' \in \Omega_{3,\delta_t}$), this shows that $\rho \subseteq \{\rho_1' \cup \rho_2'\} \cap \rho_3'$ (resp. $\rho' \subseteq \{\rho_1' \cup \rho_2'\} \cap \rho_3'$). It contradicts to our choice $\rho \in \Omega_{3,\beta_t}$. Thus, $\Omega_{3,\beta_t} \cap \Omega_{i,\beta_t} = \emptyset$ and $\Omega_{3,\delta_t} \cap \Omega_{i,\delta_t} = \emptyset$, $i = 1, 2$. Further, let $|\{\rho_1' \cup \rho_2'\} \cap \rho_3'| = k_{1,2,3}$. Then

$$|\Omega_{3,\beta_t}| = \binom{C_t}{\beta_t} - \binom{k_{1,2,3}}{\beta_t} \text{ and } |\Omega_{3,\delta_t}| = \binom{C_t}{\delta_t} - \binom{k_{1,2,3}}{\delta_t}.$$

Choose an error pattern $\rho_4 \in R_t(\beta_t)\backslash\cup_{i=1}^{3}\Omega_{i,\beta_t}$, and extend ρ_4 to an error pattern ρ_4' with $\rho_4 \subseteq \rho_4'$ and $|\rho_4'| = rank_t(\rho_4') = C_t$. Define two sets similarly:

$$\Omega_{4,\beta_t} = \{\rho \subseteq \rho_4' : \rho \in R_t(\beta_t), \rho \nsubseteq \{\cup_{i=1}^{3}\rho_i'\} \cap \rho_4'\},$$

and

$$\Omega_{4,\delta_t} = \{\rho' \subseteq \rho_4' : \rho' \in R_t(\delta_t), \rho' \nsubseteq \{\cup_{i=1}^{3}\rho_i'\} \cap \rho_4'\}.$$

For all $\rho \in \Omega_{4,\beta_t}$ and $\rho' \in \Omega_{4,\delta_t}$, $\rho \notin \cup_{i=1}^{3}\Omega_{i,\beta_t}$ and $\rho' \notin \cup_{i=1}^{3}\Omega_{i,\delta_t}$. Assume the contrary, i.e., $\rho \in \cup_{i=1}^{3}\Omega_{i,\beta_t}$, which implies that $\rho \subseteq \{\rho_1' \cup \rho_2' \cup \rho_3'\} \cap \rho_4'$. It is a contradiction. Similarly, we have $\rho' \notin \cup_{i=1}^{3}\Omega_{i,\delta_t}$ for all $\rho' \in \Omega_{4,\delta_t}$. That is, $\Omega_{4,\beta_t} \cap \Omega_{i,\beta_t} = \emptyset$ and $\Omega_{4,\delta_t} \cap \Omega_{i,\delta_t} = \emptyset$, $i = 1, 2, 3$. Let $|\{\cup_{i=1}^{3}\rho_i'\} \cap \rho_4'| = k_{1,2,3,4}$. It follows that

$$|\Omega_{4,\beta_t}| = \binom{C_t}{\beta_t} - \binom{k_{1,2,3,4}}{\beta_t}, \ |\Omega_{4,\delta_t}| = \binom{C_t}{\delta_t} - \binom{k_{1,2,3,4}}{\delta_t}.$$

We continue this procedure until we cannot choose a new error pattern $\rho \in R_t(\beta_t)$. Since $|R_t(\beta_t)|$ is finite, this procedure will stop at some step. Without loss of generality, assume that the procedure stops at the mth step. That is, $R_t(\beta_t) = \cup_{i=1}^{m}\Omega_{i,\beta_t}$. Together with what we have proved above, $\Omega_{i,\beta_t} \cap \Omega_{j,\beta_t} = \emptyset$ for all i, j satisfying $i \neq j$ $(1 \leq i, j \leq m)$. This implies that

$$|R_t(\beta_t)| = \sum_{i=1}^{m} |\Omega_{i,\beta_t}| = \sum_{i=1}^{m} \left[\binom{C_t}{\beta_t} - \binom{k_{1,2,\cdots,i}}{\beta_t} \right],$$

where set $k_1 = 0$. Similarly, we also have $\Omega_{i,\delta_t} \cap \Omega_{j,\delta_t} = \emptyset$ for all i, j satisfying $i \neq j$ $(1 \leq i, j \leq m)$, and $\cup_{i=1}^{m} \Omega_{i,\delta_t} \subseteq R_t(\delta_t)$, which implies that

$$|R_t(\delta_t)| \geq \sum_{i=1}^{m} |\Omega_{i,\delta_t}| = \sum_{i=1}^{m} \left[\binom{C_t}{\delta_t} - \binom{k_{1,2,\cdots,i}}{\delta_t} \right].$$

In order to prove $|R_t(\beta_t)| \leq |R_t(\delta_t)|$, it suffices to show $|\Omega_{i,\beta_t}| \leq |\Omega_{i,\delta_t}|$, i.e., $\binom{C_t}{\beta_t} - \binom{k_{1,2,\cdots,i}}{\beta_t} \leq \binom{C_t}{\delta_t} - \binom{k_{1,2,\cdots,i}}{\delta_t}$ for each $i = 1, 2, \cdots, m$.

To simplify the notation, we omit the subscripts in the following discussion. It follows that we just need to prove

$$\binom{C}{\delta} - \binom{k}{\delta} \geq \binom{C}{\beta} - \binom{k}{\beta},$$

that is,

$$\binom{C}{\delta} - \binom{C}{\beta} \geq \binom{k}{\delta} - \binom{k}{\beta}, \qquad (11)$$

where $\beta \leq \delta \leq \lfloor \frac{C}{2} \rfloor$ and $k \leq C$.

If $k < \delta$, the inequality (11) immediately holds. Otherwise $k \geq \delta$, note that

$$\begin{aligned} &\binom{C}{\delta} - \binom{C}{\beta} \\ =&\left[\binom{C}{\delta} - \binom{C}{\delta-1} \right] + \left[\binom{C}{\delta-1} - \binom{C}{\delta-2} \right] + \cdots \\ &+ \left[\binom{C}{\beta+2} - \binom{C}{\beta+1} \right] + \left[\binom{C}{\beta+1} - \binom{C}{\beta} \right], \end{aligned}$$

and

$$\begin{aligned} &\binom{k}{\delta} - \binom{k}{\beta} \\ =&\left[\binom{k}{\delta} - \binom{k}{\delta-1} \right] + \left[\binom{k}{\delta-1} - \binom{k}{\delta-2} \right] + \cdots \\ &+ \left[\binom{k}{\beta+2} - \binom{k}{\beta+1} \right] + \left[\binom{k}{\beta+1} - \binom{k}{\beta} \right]. \end{aligned}$$

This implies that the inequality (11) holds provided that we can show

$$\binom{C}{a+1} - \binom{C}{a} \geq \binom{k}{a+1} - \binom{k}{a}$$

for any a satisfying $\beta \leq a \leq \delta - 1$. After a simple calculation, it is equivalent to prove

$$\begin{aligned} &C(C-1)\cdots(C-a+1)(C-2a-1) \\ \geq&k(k-1)\cdots(k-a+1)(k-2a-1). \qquad (12) \end{aligned}$$

It is not difficult to see that the inequality (12) holds for $k \geq \delta$. This completes the proof. ■

References

[1] R. W. Yeung and Z. Zhang, "Distributed source coding for satellite communications," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1111-1120, May 1999.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204-1216, Jul. 2000.

[3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371-381, Jul. 2003.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782-795, Oct. 2003.

[5] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. M. G. M. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973-1982, Jun. 2005.

[6] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413-4430, Oct. 2006.

[7] H. Balli, X. Yan, and Z. Zhang, "On randomized linear network codes and their error correction capabilities," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3148-3160, Jul. 2009.

[8] X. Guang and F.-W. Fu, "On Failure Probabilities of Random Linear Network Coding," to be submitted.

[9] N. Cai and R. W. Yeung, "Network coding and error correction," *in Proc. IEEE Information Theory Workshop 2002*, Bangalore, India, Oct. 2002, pp. 119-122.

[10] R. W. Yeung and N. Cai, "Network error correction, part I: Basic concepts and upper bounds," *Communications in Infomation and Systems*, vol. 6, pp. 19-36, 2006.

[11] N. Cai and R. W. Yeung, "Network error correction, part II: Lower bounds," *Communications in Infomation and Systems*, vol. 6, pp. 37-54, 2006.

[12] Z. Zhang, "Linear network error correction codes in packet networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 209-218, Jan. 2008.

[13] S. Yang, R. W. Yeung, C. K. Ngai, "Refined Coding Bounds and Code Constructions for Coherent Network Error Correction," submitted to IEEE Trans. Inf. Theory. [Online]. Available: http://arxiv.org/abs/0904.1897.

[14] S. Yang, "Coherent network error correction," Ph.D. dissertation, The Chinese University of Hong Kong, 2008.

[15] S. Yang, R. W. Yeung, and Z. Zhang, "Weight properties of network codes," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 371-383, 2008.

[16] X. Yan, H. Balli, and Z. Zhang, "Decode Network Error Correction Codes beyond Error Correction Capability," preprint.

[17] R. Koetter and F. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3579-3591, Aug. 2008.

[18] Z. Zhang, "Network error correction coding in packetized networks," *in Proc. IEEE Information Theory Workshop 2006*, Chengdu, China, Oct. 2006, pp. 433-437.

[19] R. Matsumoto, "Construction algorithm for network error-correcting codes attaining the singleton bound," *IEICE Trans. Fundamentals*, vol. E90-A, no. 9, pp. 1729-1735, Nov. 2007.

[20] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, "Network coding theory," *Foundations and Trends in Communications and Information Theory*, vol. 2, nos.4 and 5, pp. 241-381, 2005.

[21] R. W. Yeung, *Information Theory and Network Coding*. New York: Springer, 2008.

[22] Z. Zhang, *A Course of Network Coding at Nankai University*, 2009.